Photonics

edited by
M. BALKANSKI
P. LALLEMAND

GAUTHIER - VILLARS

Photonics

Photonics

edited by
M. BALKANSKI
P. LALLEMAND

GAUTHIER - VILLARS

ISBN 2-04-005023-X



EXPLANATORY NOTE

PHOTONICS : PERSPECTIVES IN 1974

==============

The development of new technologies and their rapid transfer into current industrial practice is one of the objectives of the French General Delegation for Scientific and Technological Research (D.G.R.S.T.). Optoelectronics is a field with great potential because of a rapid developments that recently occurred in semiconductor lasers, in photodetectors and in new materials suchas optical fibers with low optical losses.

In order to develop the field of optoelectronics, the D.G.R.S.T. found it necessary to obtain broad information and to accumulate a complete documentation in this field. For this purpose a group of physicists has been designated to organize meetings and discussions among specialists, not only from France, but from all countries where research in this field is at a substantial level.

The word *photonics* is used throughout by analogy with electronics, to designate all functions where a photon rather than an electron is the information vehicle. The field of photonics begins when a laser diode, for example, converts an electronic signal into a photon beam. Photonic functions replace the well known electronic functions : propagation, deflection, modulation, amplification and storage of light.

Our interest in this exploratory program was twofold; we thought it interesting to define as well as possible the functions which play a role in this new technology as well as to review the physical phenomena which form the basis of the new devices. During the past year we have called seminars where physicists and engineers have met to define and compare current performance possibilities. From these seminars and discussions, the clear conclusion is that the area which will lead to better devices is that of material characteristics. The field of materials is where the largest effort will be developed and where the largest number of research projects will begin. It is only when a high degree of control in this field is obtained that one can really hope to obtain higher performance in the field of

photonics than that of electronics. Better performance is not the only advantage of photonics. The use of new materials such as optical fibers (which substitute for metallic conductors) will eventually be of economical importance.

In order to summarize the present development of this field, the D.G.R.S.T. has sponsored and organized an international meeting whose proceedings are given in this volume. For each specific aspect of photonics an invited paper was given by one of the best specialists in the field. Several talks which were given during the 1973-1974 academic year have been added to the proceedings of the conference which was held from June 27th to 29th, 1974 in Cadarache. The set of these papers constitute a general review of the field of photonics which promises a rapid development in the coming years. Some of the papers which will be found in this volume will become historical documents ; others will serve as the basis and guides for further research and development in future years. These papers taken together give the present state of the development of photonics and certainly constitute a valuable guide for new research fellows entering this field. All the elements of photonics are described and discussed with all the details possible wuth today's knowledge.

The most important field of research in the first phase of development of optical transmission is certainly that of optical fibers and the investigations on optical losses in such fibers. It appears that optical losses in glasses available today are already lower than the losses known for the metallic conductors of general use. It remains, nevertheless, to solve the problems of connecting and coupling the different optical devices to light guides ; further research in this field will certainly affect the future technological development. The efficiency of communication by optical fibers is optimized by the existence of double-heterojunction lasers such as GaAs-GaAlAs, having a wavelength which is exactly in the region of lowest attenuation of optical fibers and which is also in the frequency region of high sensitivity of silicon avalanche photodiodes.

It remains, of course, to devote a great deal of research in order to improve the lifetime of semiconductor lasers. In addition,the technology of pulling optical fibers is far from attaining the limit of the lowest attenuation physically possible. In this case, again, technological development as well as fundamental research will probably stimulate rapid progress. Research developed in the field of photodiodes has found, today, a new justification in the use of photodiodes in the optical communications system. All this justifies the opinion that communication by optical fibers will be put into use rapidly. It nevertheless remains to solve new and numerous technological problems of great importance such as the lifetime of the equipment, the plastic coating of glass fibers and their cabling, the problem of splicing and that of installation and maintenance.

Although the physical principles of the modulation and deflection of optical beams are known, it is necessary to make further progress in the field of electrooptic materials. This certainly will be a field of extremely useful research for the future.

As far as amplification is concerned, there is not yet a clear cut opinion probably due to the general lack of experience in systems of this type. The specialists do not agree on the importance to be given to the function of application in optical transmissions. Nevertheless, it is possible that it will become necessary to face this problem in the near future. It will then be necessary for physicists to find methods of optical amplification and the means to put these methods into practical use.

The first step towards achieving optoelectrical systems has been made in the field of integrated optics where spectacular progress has been realized. However we still have a long way to go in order to achieve a level of integration comparable to that known today in electronics.

While the possibilities of optical communication are easily forseeable today, it may be a little different for optoelectronic systems which are further away for applications in information storage. Participants at this conference have nevertheless stretched their imagination and discussed the future of this application.

This document is almost a birth certificate for a new technological field which will certainly be developed in the near future. It gives the present status of development and we hope will also be an encouragement for further research. In organizing the coordination in this field the D.G.R.S.T. has designated the role which it will play in the further developement of this technology.

M. BALKANSKI

CONTENTS

LIST OF AUTHORS

P. AIGRAIN
Thomson C.S.F. Bd Haussmann - 75008 PARIS France

F.T. ARECCHI
Università di Pavia and Cise MILANO - Italy

J. BEIMAS
Laboratoire de Marcoussis
Département Recherches Physique de Base
Section Composants Electroniques CGE
91460 MARCOUSSIS - France

B. DESORMIERE et G. HEPNER
Laboratoire Central de Recherche
Thomson C.S.F.
Domaine de Corbeville
91401 ORSAY - France

J. ERNEST
Laboratoire de Marcoussis
Département Recherches Physique de Base
Section Composants Electroniques CGE
91460 MARCOUSSIS - France

A.M. GLASS
Bell Laboratories
Murray Hill
NEW Jersey 07974 U.S.A.

J.P. HAZAN
Laboratoire d'Electronique et de Physique Appliquée
94450 Limeil-Brévannes - France

I.P. KAMINOW
Bell Laboratories Crawford Hill Laboratories
Holmdel , N.J. 07733 U.S.A.

H. KOGELNIK
Bell Laboratories
Holmdel N.J. 07733 U.S.A.

H. KRESSEL
R.C.A. Laboratories
Princeton - New Jersey 08540 U.S.A.

H. KURZ
Philips Forschungs Laboratorium Hamburg GmbH
HAMBURG 54 Germany

H. LE GALL
Laboratoire de Magnétisme et d'Optique
des Solides C.N.R.S. 92190 MEUDON-BELLEVUE -France

F. MICHERON
Thomson - C.S.F.
91401 ORSAY - France

D.B. OSTROWSKY
Thomson-C.S.F.
Domaine de Corbeville
B.P. n° 10 91401 ORSAY - France

D.B. OSTROWSKY and J. SÈVIN
Thomson - C.S.F.
B.P. n°10 91401 ORSAY - France

J.I. PANKOVE
R.C.A. Laboratories Princeton
New Jersey 08540 U.S.A.

F.H. RAYMOND
Thomson C.S.F.
Boulevard Haussman 75008 PARIS

E. SPITZ
Thomson C.S.F.
Domaine de Corbeville
91401 ORSAY - France

M. TREHEUX, A. COZANNET, R. BOUILLIE
C.N.E.T. 22301 LANNION - France

T. UCHIDA
Nippon Electric Co., Ltd
Japan

M. BALKANSKI
Université de Paris VI Tour 13
4, Place Jussieu 75005 PARIS

P. LALLEMAND
Ecole Normale Supérieure
Laboratoire de Physique
24, Rue Lhomond 75231 Paris Cedex 05 - France

SEMICONDUCTOR ELECTROLUMINESCENT DIODES
FOR FIBER-OPTICAL COMMUNICATIONS

H. Kressel

R.C.A. Laboratories

Princeton, New Jersey 08540 U.S.A.

I. INTRODUCTION

There has been extensive research devoted to the development of practical electroluminescent diode sources suitable for optical communication systems, particularly following the successful realization of efficient room temperature laser diodes by incorporating (AlGa)As-GaAs heterojunctions. [1,4] This effort has received additional impetus from the development of low-loss (<20 db/km) glass fibers which have several low absorption spectral regions, including the 8000 - 8800 Å [5] ideally suited for (AlGa)As heterojunction lasers or incoherent radiation sources.

In addition to having the appropriate emission wavelength, diodes useful in fiber-optical communications should have a number of desirable properties in order to fully realize the potential of this new technology.

1) The diodes should be capable of modulation at rates in excess of 100 MHz for high data rate transmission (although applications exist for much lower modulation rates).
2) The spectral width should be narrow in order to minimize the pulse dispersion in the fibers which limits the bandwidth.
3) The average power needed is typically on the order of milliwatts, although higher power values are needed for very long continuous fiber links, or if higher loss fibers are used.
4) The radiance of the source should be as high as possible for effective coupling into the lowest loss fibers with their small numerical apertures (∿0.14). This means that the beam spread of the diode emission should be minimized.
5) The diode life should be in the tens of thousands of hours, although some systems in intermittent use require more modest lifetimes.

A laser diode capable of reliable CW operation at room temperature constitutes, in principle, the ideal radiation source for fiber optical communications, since modulation rates of ∿1 GHz have

been achieved. However, the incoherent light-emitting diode (LED) technology, also based on the use of heterojunction structures, is simpler. Furthermore, because of their lower current density of operation, LEDs are easier to use than lasers. Thus, LEDs fulfill a number of potential systems requirements for relatively short distance, moderate bandwidth, optical links. Various relevant aspects of laser diodes and LEDs are reviewed here, with emphasis on general principles rather than engineering design aspects. For introductory material dealing with basic device physics we refer to earlier reviews. [6,7] In Section II, low threshold current density laser diodes are discussed. Section III considers high speed, light-emitting diodes, and in Section IV the subject of diode degradation phenomena relevant to both CW lasers and light-emitting diodes is reviewed.

▪ II. LOW THRESHOLD LASER DIODES

A. Structures

Figure 1 shows a five-layer vertical diode structure which is the generalized GaAs injection laser ; [8] presently used diodes are variations of the structure which contain either some or all of the layers shown. Basically, the active region of the laser consists of a radiative recombination region 3 of width d_3 , the " gain region," and a waveguiding region which is either equal to or thicker than the gain region. In Fig.1 the nominal waveguide region is of thickness $w = d_2 + d_3 + d_4$ with the higher bandgap regions 1 and 5 providing the walls of the waveguide because of their reduced index of refraction at the GaAs lasing photon energy. The difference in the Al content between i-numbered layers determines the refractive index difference Δn, which at the typical 300K GaAs lasing emission wavelength of $\sim$ 9000 Å , increases approximately linearly with the bandgap energy difference ΔE_g. For example , for $\Delta E_g = 0.4$ eV, $\Delta n \approx 0.2$ or $\Delta n/n \cong 0.2/3.6 \approx 6\%$ [9]. Carrier confinement requires a potential barrier height of a few kT. At 300 K, a bandgap energy difference of $\gtrsim 0.1$ eV is adequate.

Various heterojunction laser structures are obtained by using selected layers in Fig.1. If $d_4 = d_2 = 0$, we obtain the single heterojunction laser. [1,2] . If $d_2 = d_3 = 0$ and $x(5) = x(1)$, we obtain the symmetrical double heterojunction (DH) laser. [3] With $d_2 = 0$ and $x(4) = 0$, the "large optical cavity" (LOC) laser diode is obtained. [10] A four-heterojunction laser is obtained when all the regions are used. [11] Each of these structures is capable of a wide range of properties depending on the width of the various regions, their Al content, symmetry and doping in the GaAs recombination region.

At a given temperature, the laser threshold current density, J_{th}, is a function of the optical waveguide thickness w, the width of the recombination region d_3 , the doping level, the facet reflectivity R and the length of the laser L. The gain coefficient required to reach threshold, g_{th}, is :

$$g_{th} = \bar{\alpha} + \frac{1}{L} \ln \frac{1}{R} \qquad (1)$$

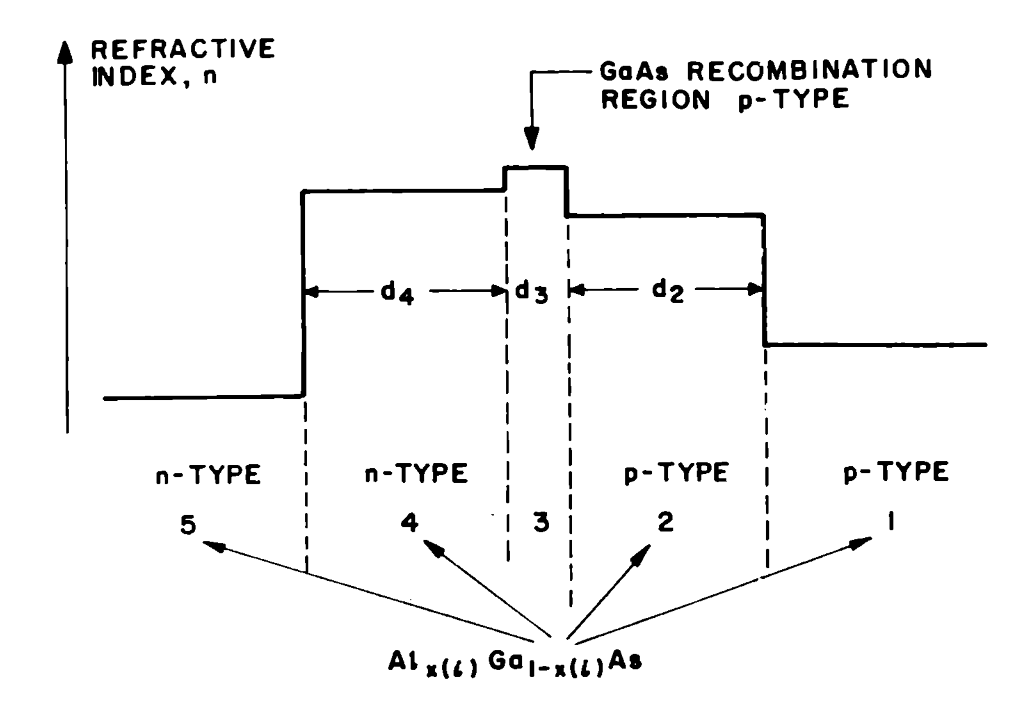


Figure 1 :
Cross-section of generalized laser structure used in analysis of transverse modes, showing the refractive index at the lasing frequency. The radiative recombination occurs in region 3, while regions 2,3, and 4 constitute the nominal waveguide. (Ref 8).

where $\bar{\alpha}$ is the effective absorption coefficient. The major contribution to $\bar{\alpha}$ is due to free carrier absorption ($\alpha \cong 10\ cm^{-1}$) in well designed heterojunction lasers.

The gain coefficient for a material of given doping is a function of the injected carrier density n in the recombination region (at threshold $n \propto J_{th}/d_3$), of the form $g \propto n^b$ where $b \geq 1$, and of the fraction, Γ , of the radiation which propagates through the recombination region

$$\Gamma \cong d_3/w \qquad (2)$$

Therefore,

$$g_{th} \propto \Gamma n^b \propto \Gamma (J_{th}/d_3)^b \qquad (3)$$

(We neglect a coupling factor which takes into account the position of the gain region within the waveguide region).

Combining Eqs. 1-3, we obtain an expression for J_{th},

$$J_{th} \propto w^{1/b} d_3^{(1-\frac{1}{b})} [\bar{\alpha} + \frac{1}{L} \ln \frac{1}{R}]^{\frac{1}{b}} \tag{4}$$

It is evident from Eq. 4 that if b = 1, then J_{th} increases linearly with the width of the waveguide region 2, assuming $\bar{\alpha}$ constant. However if b > 1, then for constant w, it is possible to reduce j_{th} by reducing the recombination region width d_3.

The differential quantum efficiency η_{ext} is determined by the Fabry-Perot cavity properties and the value of $\bar{\alpha}$,

$$\eta_{ext} \approx \eta_i \frac{\ln \frac{1}{R}}{\bar{\alpha}L + \ln \frac{1}{R}} \tag{5}$$

where η_i, the internal quantum efficiency, is typically 0.6 - 0.7 in good quality devices at room temperature.

The doping level and degree of compensation determine the dependence of the gain coefficient on injected carrier density, a relationship which has been studied both theoretically and experimentally. At room temperature, the value of b is theoretically predicted to increase from ∿1 to ∿3 as the material purity is increased. [12, 13] Figure 2 shows a plot of the gain coefficient vs. threshold current density calculated [13] for lightly doped material (with and without momentum conservation for band-to-band recombination) and for heavily doped material. Figure 2 also shows that good agreement of experiment [14] with theory is obtained at the extremes of high and low doping. $b \approx 1$ in closely compensated p-type GaAs : Si ($N_D + N_A = 10^{19}$ cm^{-3}, $p \approx 5 \times 10^{17}$ cm^{-3}), while $b \cong 3$ (for low gain values) in a lightly doped n-type recombination region ($N_A + N_D = 2.5 \times 10^{16}$ cm^{-3}), $n \cong 2 \times 10^{16}$ cm^{-3}).

The higher gain coefficient obtained from a given injected carrier density with highly doped material makes Si-doping particularly advantageous to obtain low J_{th} DH lasers. In fact, a typical J_{th} value for such Si-doped DH lasers is ∿ 3-3.5 KA cm^{-2} per μm of active region width compared to ∿5 KA cm^{-2} μm^{-1} for similar lasers with undoped active regions.

The lowest room temperature J_{th} value of 575 A/cm^2 has been obtained [15] with a four-heterojunction laser diode where the recombination region was exceptionally thin d_3 =0.04 μm and and w=0.44 μm (Fabry-Perot cavity length L= 525 μm). This J_{th} value, substantially lower than obtained with the DH configuration with the same w value, is a consequence of Eq. 4 with b > 1. However, the control of such narrow d_3 regions needed to obtain J_{th} < 1000 A/cm^2 is still difficult with the present technology.

B. Beam Patterns

Extensive effort has been devoted to a study of the modal properties of laser diodes because of the simplification in the

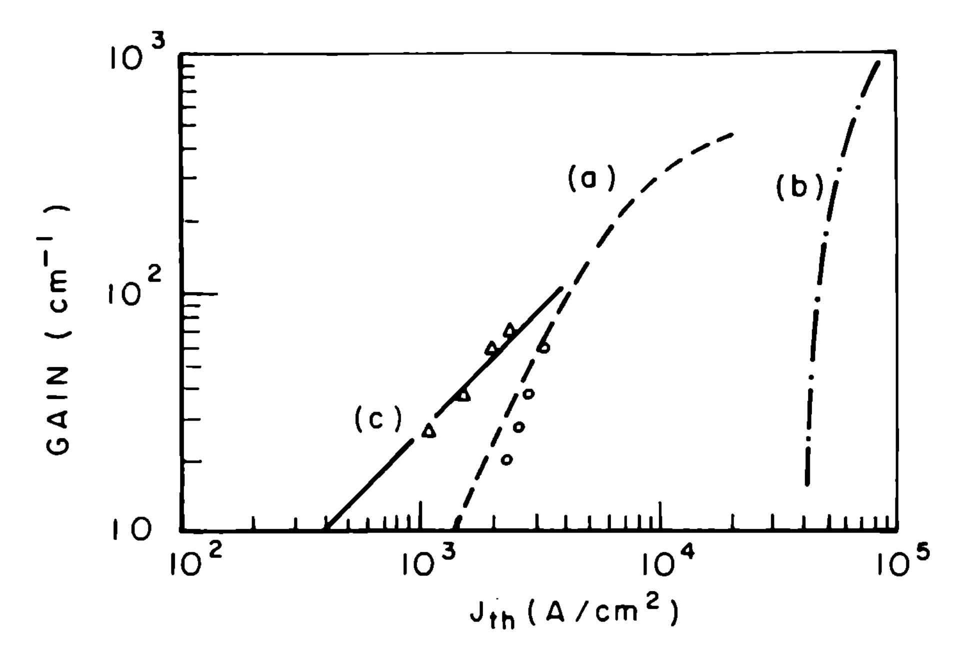


Figure 2 :
Gain per unit length versus threshold current density for GaAs injection lasers : (a) Theoretical, [13] parabolic bands, k-selection, (b) theoretical, [13] parabolic bands, no k-selection, (c) theoretical, [13] exponential tails, no k-selection.
-O- Experimental p^+nn^+: [14] $N_D+N_A=2.5\times10^{16}cm^{-3}$ $n=2\times10^{16}cm^{-3}$.
-Δ- Experimental [14] p^+pn^+ : $N_D+N_A=10^{19}cm^{-3}$, $p\sim5\times10^{17}$ cm^{-3}.

optical collimating system when the far-field beam profile is reduced. Arbitrarily increasing the width of the waveguiding region not only increases the threshold current density but also results in the propagation of high order transverse modes (i.e., several maxima in the electro-magnetic field intensity) and consequently rabbit-ear far-field patterns. This is illustrated in Fig. 3.[16] Conversely, decreasing the heterojunction spacing decreases the threshold current density (if the refractive index steps are sufficiently large for full radiation confinement), but at the expense of a very broad beam which can reach ~80°. Note that in almost all laser diode devices, the far-field pattern *parallel* to the junction plane has a half power angular spread of 8-15°.

A practical structure which offers a useful compromise between simplicity of fabrication, moderate beam divergence and sufficiently low J_{th} for CW operation is obtained by the use of small

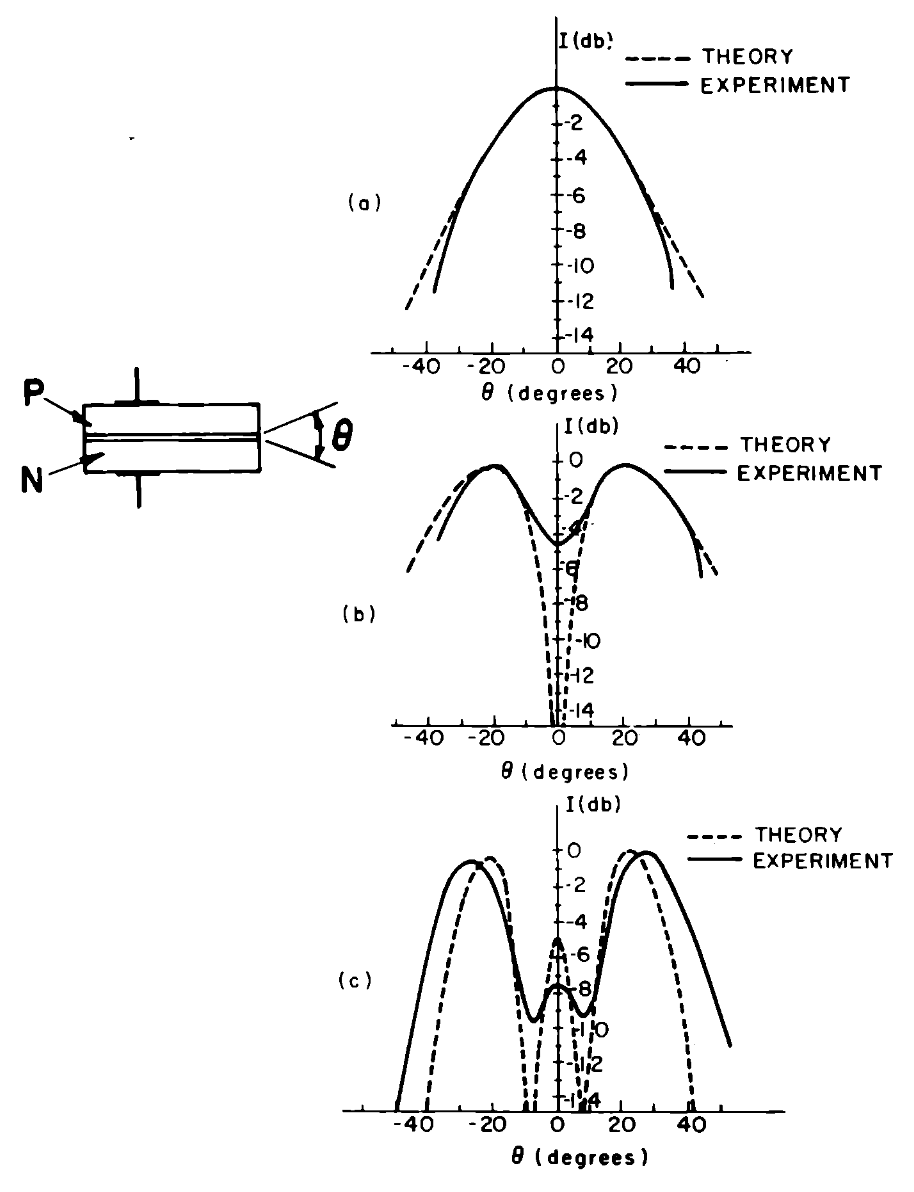


Figure 3 :
Comparison of the experimental and theoretical far-field patterns perpendicular to the junction for double heterojunction lasers with varying thickness of optical cavity d :
(a) 0.7 µm, fundamental mode only ;
(b) I.I µm, second-order mode only ;
(c) 2.8 µm, third-order mode with small admixture of the fourth-order mode. (Ref I6).

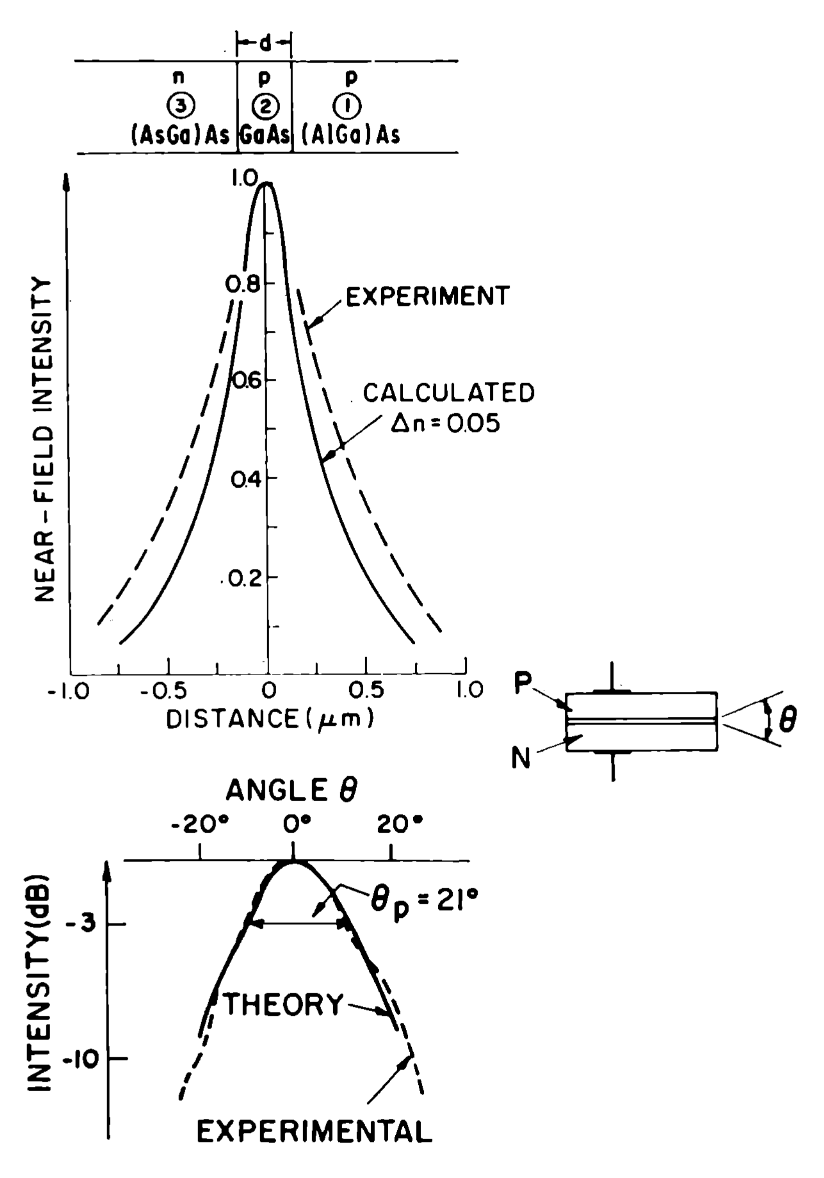


Figure 4 :

Data illustrating the effect of partial radiation confinement in specially constructed double heterojunction diode for reduced radiated beam spread. (a) Near-field intensity distribution compared to the theoretically calculated value ; (b) far-field radiation pattern in the direction perpendicular to the junction plane showing a beam spread at the half-intensity point of ∿21°. (Ref 17).

heterojunction spacings (0.2 - 0.3 μm) with the symmetrical bandgap energy step adjusted to give only partial wave confinement but complete carrier confinement. [17] Only the fundamental mode can then be excited, and since the GaAs radiation spreads into low loss passive (AlGa)As, the optical losses are low. Figure 4 shows the near-and far-field radiation pattern of such a device. The intensity distribution at the emitting facet perpendicular to the p-n junction is seen to extend beyond the heterojunction boundaries as expected theoretically. The threshold current density of such lasers is about 1400-3000 A/cm^2 with a beam divergence of 21-30° and a differential quantum efficiency of 30-50%.

The above design concept is, of course, applicable also to structures where the recombination region consists of a desired (AlGa)As composition, with outer regions appropriately adjusted to give the required bandgap energy step. Although the fabrication technology increases in difficulty with increasing Al content, the room temperature threshold current density and efficiency are relatively constant as the emission wavelength is chosen to be between 8000 and 9000 Å, making it possible to reproducibly fabricate CW laser diodes within that spectral range.

C. Diode Construction

The devices are epitaxially grown from the liquid phase as reviewed in detail elsewhere. [18,19] Epitaxial growth of AlAs-GaAs alloys is greatly simplified by the almost perfect match in lattice parameter between these two compounds, but multiple layer growth of heterojunction structures involves growing with excellent control adjacent layers with widely varying compositions and dopings. The interfaces between these layers must be flat ; there must be no contamination from one growth solution to the next ; the layer thickness must be precisely controlled with sub-micron tolerances; and the final surface must be free of any solution when the wafer is withdrawn from the growing apparatus.

Basically, the growth technique used consists of sliding a GaAs substrate sequentially into bins containing various solutions. Of the many possible designs of the growth apparatus, variations of the linear multiple-bin graphite boat [20] have proven the most popular. In a model shown in Fig. 5(a), a GaAs source wafer, usually polycrystalline, precedes the substrate wafer into each bin and assures saturation of the solution before growth on the substrate is initiated. [21] Figure 5(b) is an improvement over this design in that each solution has its own source wafer, leading to improved layer control. In a well fabricated wafer, there are gradual steps in this surface with a height of only 0.1 μm or less ; thus, the wafer can be processed for ohmic contacting without the need of a polishing step. This is especially important when the active layers of the structure are within a few micrometers of the surface as they are in the CW laser.

Room temperature CW operation is feasible when the threshold current and heat dissipation are such that the temperature rise is limited to the range where the device is capable of reliable operation. In general $J_{th} \lesssim 3000$ A/cm^2 is required with

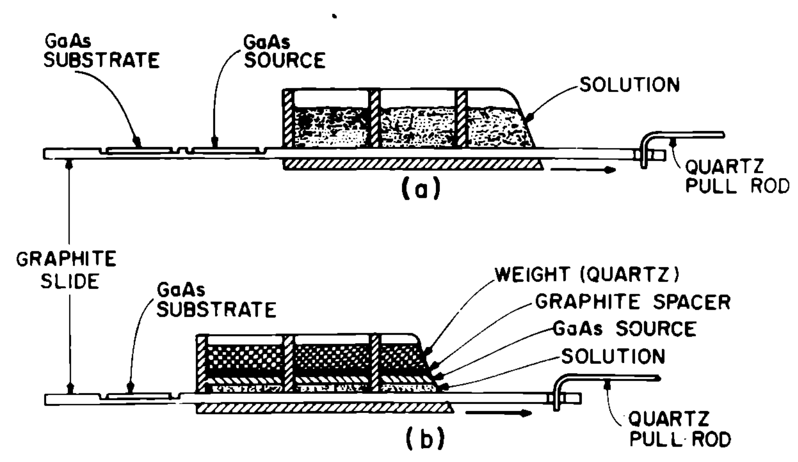


Figure 5 :
Schematic illustration of growth boat. In (a) the saturation of the solutions is completed by the source wafer preceding the substrate wafer while in (b) each solution is saturated by its own source wafer (Refs. I9 and 2I).

"broad area" diodes, i.e., where the whole chip is active and the distance between the active region and the heat sink $\lesssim$ 5 μm. Significant power levels have been obtained, even with a Cu rather than a diamond heat sink. For example, with a threshold current density of ~2000 A/cm^2 , ~120 mW have been obtained CW from one diode at a current density of 5000 A/cm^2 (7% power conversion efficiency). [17]

Since only a few milliwatts of CW power are generally adequate, the "stripe-contact" rather than broad area laser structures are used. In the planar stripe devices, the active region is imbedded in a much wider non-active area. There are several important advantages to the stripe contact : (1) The radiation is emitted from a small area which simplifies coupling to a fiber. (2) The operating current is low, usually under 0.5 A. (3) The power dissipation is increased since the heat generating emission region is imbedded in a cooler, non-active semiconductor body. (4) The small active area improves the possibility of obtaining a defect-free area which is helpful in obtaining reliable operation (see below, Section IV).

There are several techniques for making stripe-contact diodes. One widely used method consists of a planar structure where the diode is defined by a narrow metal contact, Fig. 6.[22] In another structure, a mesa is etched and the metallization extends over the mesa edge with an insulating layer to prevent shorting of the p-n junction. [23] Planar stripe lasers can also be made by selective proton bombardment [24] and various selective diffusion [25] methods to introduce high conductivity regions for the active diode area surrounded on both sides by high resistivity blocking regions.

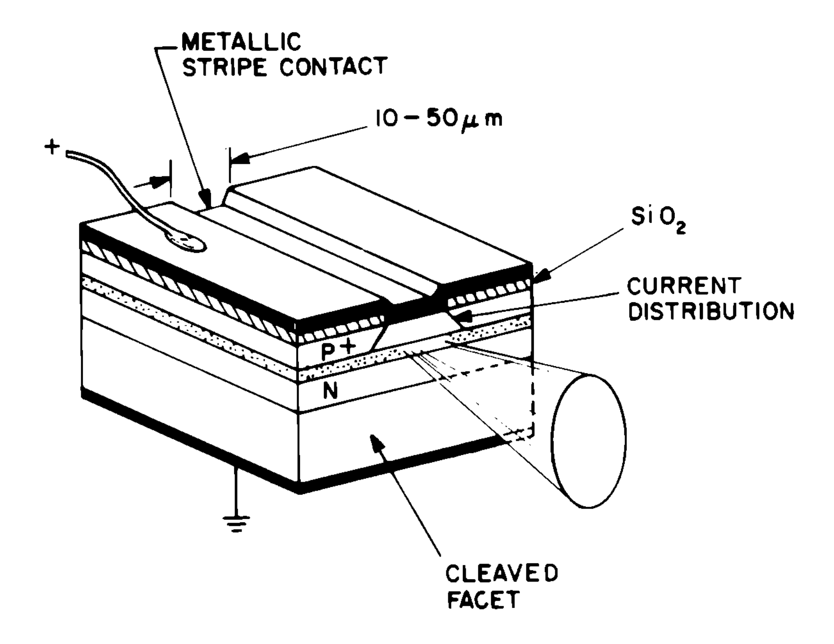


Figure 6 :
Schematic of stripe-contact laser diode using oxide insulation designed for obtaining lateral current confinement. The high resistance regions adjoining ohmic contact stripe can be also obtained by several methods discussed in the text. The diode is usually mounted with the stripe contact adjoining the heat sink to minimize the diode thermal resistance.

The planar stripe width cannot be made arbitrarily narrow without degrading the diode characteristics because of current and radiation spreading outside of the contact area. Current spreading is limited with proton bombardment or selective diffusion because the resistance of the material surrounding the active area defined by the stripe is very high, but the problem of radiation losses remains. [26,27] In practice, a stripe width of 10-15 μm offers an advantageous compromise offering low operating currents and adequate power for fiber systems. Even with this stripe width, however, the threshold current density is several times higher than that of broad area diodes fabricated from the same material. Because of the presence of bounding walls, the mesa-stripe configuration has the best current confinement and lower diffraction loss, and hence the lowest stripe-contact threshold current density has been obtained -900 A/cm^2 .

It is evident from the above summary, that many approaches exist for the formation of small-area diodes. The choice of techno-

logy may affect the operating lifetime of the diodes, although no published information is available at this time. In general, the oxide stripe technology is the simplest having relatively few operations in the production of a useful laser, but the mesa-stripe device offers the lowest operating current at the expense, however, of a more complex fabrication process.

To illustrate the actual cross section of a stripe device, we show in Fig. 7 a diode formed by selective diffusion. The semicircular region is the low resistance Zn diffused material. The diffusion front is only ~ 1 μm from the recombination region, thus providing current confinement to a narrow region ~ 20 μm wide in the diode shown. The active region is ~ 4 μm from the Cu heat sink to which it is bonded with a thin In contact. Figure 8 shows the power

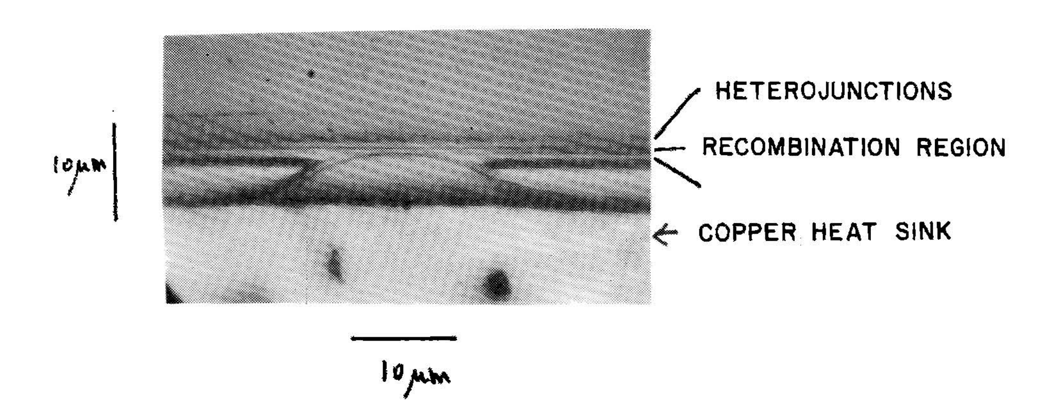


Figure 7 :
Microphotograph showing a cross-section of CW laser diode mounted on a Cu heat sink. The low resistance Zn-diffused region appears as a semicircular region.

output as a function of dc current for such a state-of-the-art diffused stripe-contact diode at room temperature. Since 5-10 mW is adequate for the anticipated application, a diode current of about 300 mA is sufficient . The emitted radiation of CW lasers is typically concentrated in a 20-30 Å spectral width as shown in Fig. 9. Some units, however, do show substantial power emission in only one or two longitudinal modes. The reason for the narrow emission occasionally seen may be due to exceptional crystalline perfection in these diodes.

With regard to modulation, GaAs double heterojunction lasers have been modulated at ~1 GHz by maintaining a dc bias to eliminate delay effects. [28]

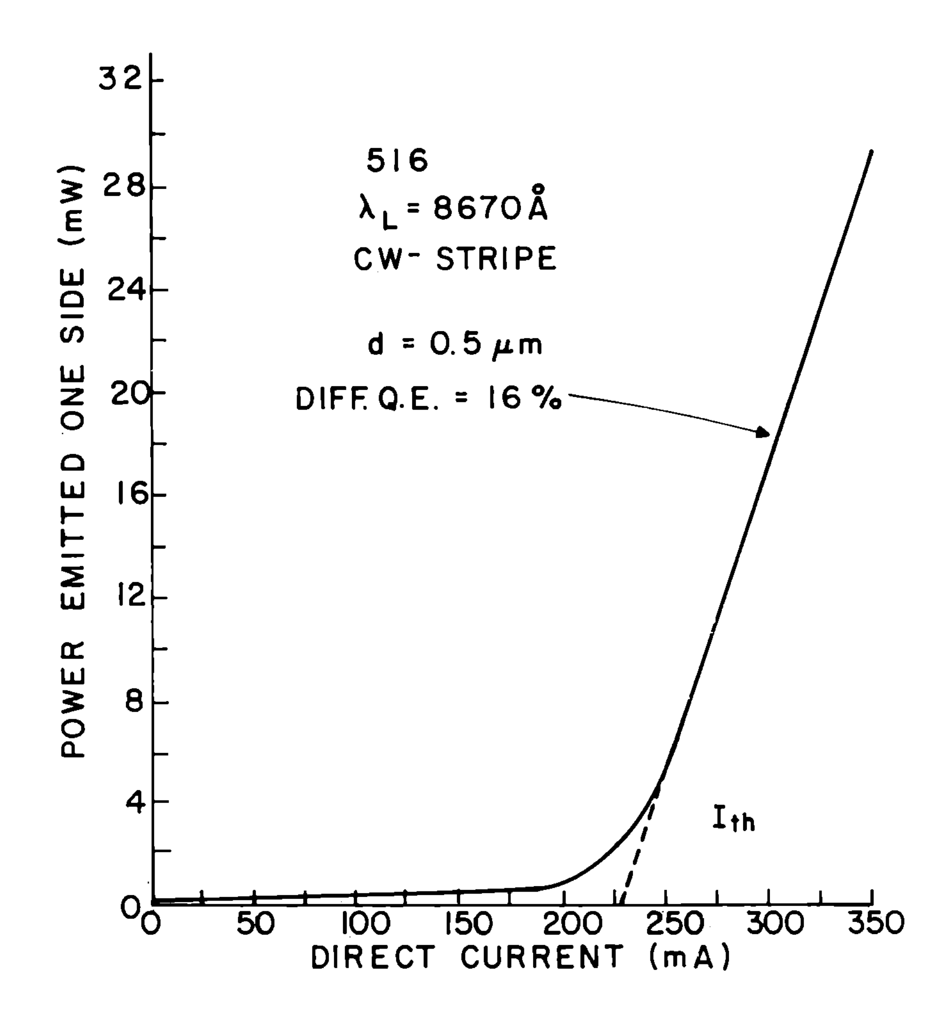


Figure 8 :
Power emitted from one facet of a state-of-the-art laser diode operating CW at room temperature. The stripe-contact region was formed by selective diffusion, as shown in Fig. 7.

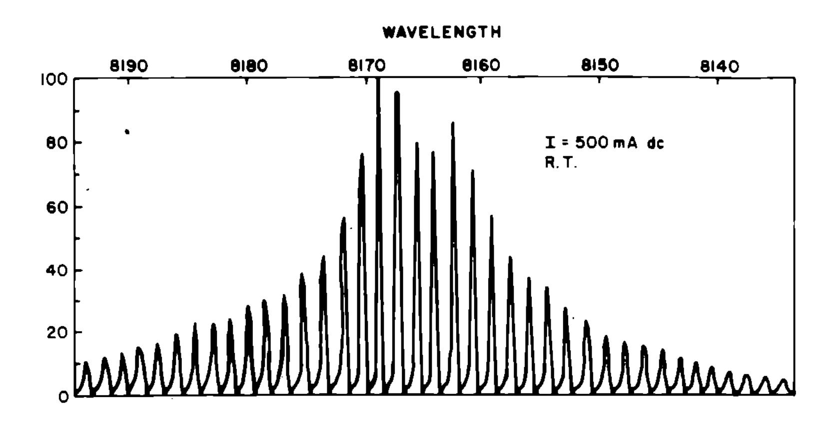


Figure 9 :
Emission spectra of typical stripe-contact diode operating CW at room temperature. The active region of this diode is of (AlGa)As.

III. HIGH SPEED LIGHT-EMITTING DIODES

The use of heterojunction structures is advantageous for the construction of high radiance LEDs because it is possible to combine an appropriately doped recombination region with surrounding material having a higher bandgap energy, hence low absorption of the emitted radiation. Furthermore, some of the advantages of heterojunction structures for lasers are also useful in spontaneous emission such as the carrier confinement which leads to high internal efficiencies and some waveguiding which leads to improved edge emission efficiency.

The minority carrier lifetime in the recombination region is a function of the carrier concentration. Under conditions where the injected carrier density is low relative to the background carrier concentration the radiative lifetime $\tau_r \propto p^{-1}$ or n^{-1} for p-type or n-type material, respectively. The radiative efficiency is a function of the ratio of the radiative to the nonradiative lifetime,

$$\eta_i = \frac{1}{1 + \frac{\tau_r}{\tau_{nr}}} \qquad (6)$$

where τ_{nr} is the nonradiative lifetime.

To increase the diode speed without sacrifice in η_i requires decreasing τ_r, by increasing the doping level, but without also

unduly decreasing τ_{nr}. It is a common observation in GaAs and related compounds that non-radiative centers are formed when the dopant concentration approaches the solubility limit under the given growth conditions. [18]

Germanium is a particularly convenient acceptor dopant in GaAs-(AlGa)As prepared by liquid-phase epitaxy because of its relatively low vapor pressure which minimizes cross-contamination of the various layers grown in a multiple-bin boat (in contrast to Zn, for example, which has a much higher vapor pressure and hence is more troublesome to use in LPE). The doping level with Ge can be increased to about $10^{19}cm^{-3}$ without sacrifice in the radiative efficiency, yielding minority carrier lifetimes of < 5 nsec. Figure 10 shows the relative efficiency of GaAs : Ge as a function of hole concentration as measured by

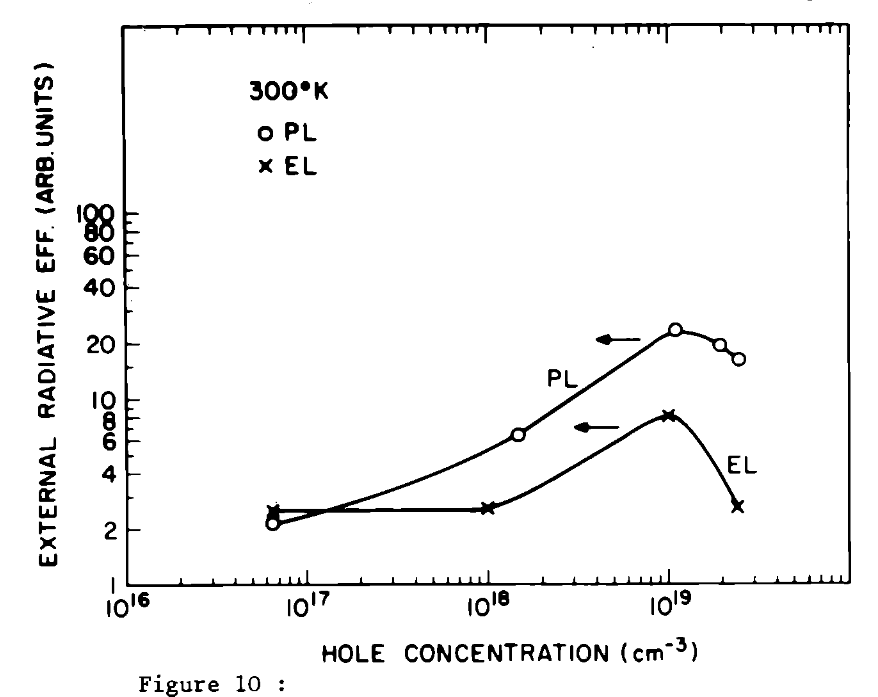


Figure 10 :
Relative photoluminescence efficiency (PL) of GaAs : Ge as a function of hole concentration, and electroluminescence (EL) of n^+p homojunction diodes fabricated using this material in the p-region of the diodes. (Ref 29).

photoluminescence and electroluminescence in p-n junctions. [29] The efficiency is seen to peak at a hole concentration of about $10^{19} cm^{-3}$. Double heterojunction diodes highly doped with Ge have rise and fall times under 5 nsec.

A short radiative decay time is reflected in the output modulation capability at high frequency. Figure 11 shows that the DH

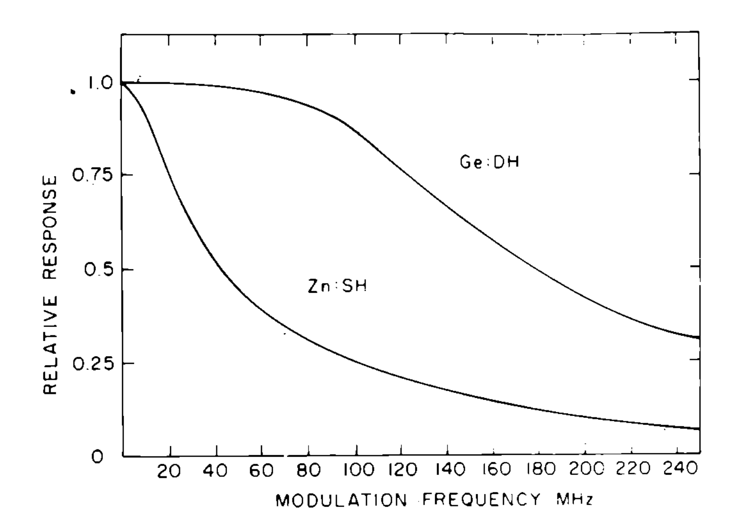


Figure 11 :
Relative optical response as a function of frequency of a double heterojunction diode with heavily Ge-doped recombination region compared to that of a single heterojunction diode. (Ref 30).

diode doped with Ge is capable of being modulated (to the 50% power output point) to 180 MHz. [30] In comparison, we show similar data for a Zn-doped single heterojunction diode with its lower modulation capability. The radiative decay time of the particular SH diode shown was ∿15 nsec.

The noise emission from GaAs homojunction and heterojunction LEDs has been measured. It was concluded that above 5 kHz, the homojunction LED tested could be considered an ideal radiation source . [31] In the case of heterojunction structures, it was concluded that poor ohmic contacts provide the dominant noise source, and that with good contacts, the measured 1/f noise contributed by the diodes would be well below the shot noise in a practical system. [32]

Source radiance is important for effective coupling of the LED emission into fibers, but the best diode configurations depend on the fiber size and numerical aperture and on whether a single fiber or a fiber bundle is used. The basic problem in the use of LEDs with optical fibers is obtaining an adequate power level launched into the fiber or fiber bundle. Because of the isotropic nature of the spontaneous emission process and the high refractive index of GaAs, only a few percent of the internally generated radiation can be emitted in a flat diode configuration. Since the surface emission is isotropic (or very nearly so), the power collected with low numerical aperture, low loss Corning fibers (N.A. = 0.14) is well under 1 mW at practical current densities from conventional surface emission diodes ($\lesssim$1000 A/cm^2) .

Higher current densities are possible with surface emission structures by the use of an etched well configuration shown in Fig 12(a) where the active diode area is small and placed close to a heat sink. [33]

A parabola configuration shown in Fig 12 (b) is useful for radiation coupling into fiber bundles. The diameter of the parabola is adjusted for the bundle diameter. By combining the emission from four

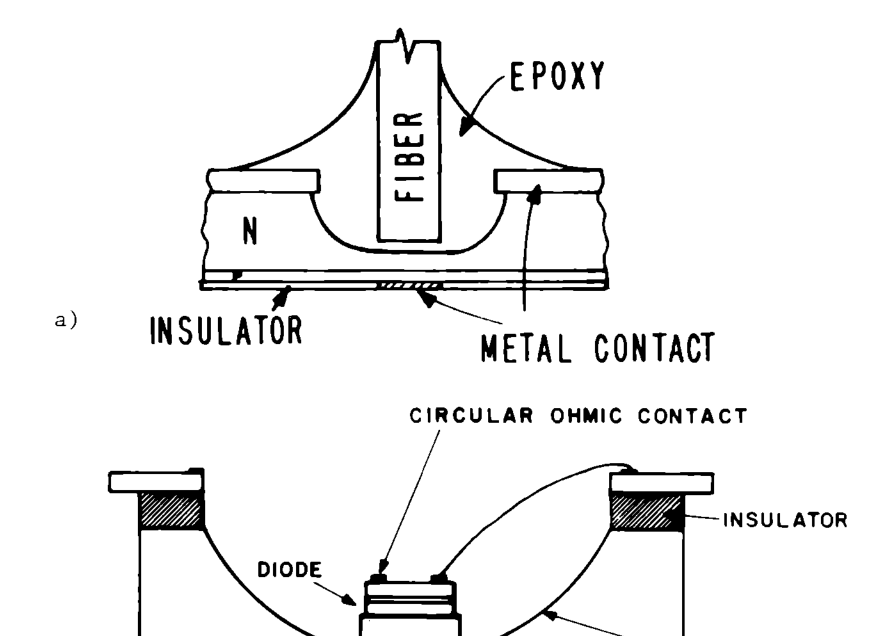


Figure 12
(a) Etched-well surface emission LED.[33]
(b) Parabola assembly of an LED combining edge and surface emission collection.

diode edges as well as some surface emission, the highest possible power emission for a given diode current density is obtained.

The edge emission utilization is attractive because of its simplicity and the fact that the radiation emitted from the edge is directional. [30,34] The emission can be imaged into the fiber edge using appropriate lenses or even butted against the fiber. Figure 13 shows examples of the power emitted as measured by placing a Si detector next to the diode and either a low loss single fiber (N.A≈.14) or a bundle of high loss fibers (N.A. ≈ 0.66). The data were obtained [35] both below and above CW laser threshold which is indicated

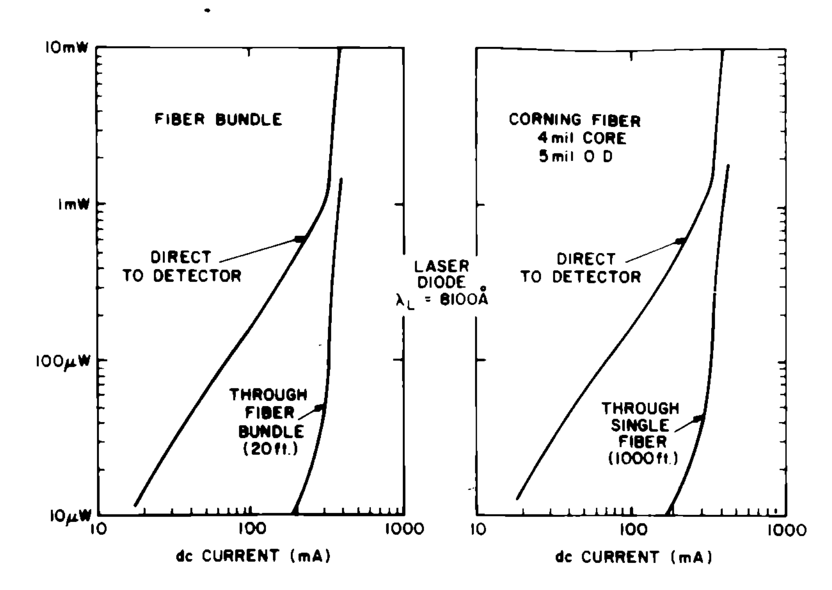


Figure 13 :
(a) Data [35] illustrating the effect of butting an edge-emitting diode against a fiber bundle containing I80 fibers with a numerical aperture of 0.66. The diameter bundle was 45 mils. The knee in the curve of power vs. current indicates the onset of CW lasing threshold.
(b) Similar data obtained using a single low loss fiber with a numerical aperture of 0.I6. The (AlGa)As laser diodes are described in Ref 36.

by the knee in the curve of emitted power vs. diode current. The diode used had a stripe contact and an emitting region of about 20 μm. [36] It is evident that tens of microwatts of power can be launched in the incoherent mode of operation, while milliwatts are launched above threshold where both the diode efficiency and directionality improve sharply.

IV. DIODE RELIABILITY

The determination of the factors which affect the stability of III-V compound electroluminescent diodes has been unusually difficult because of nonreproducible effects due to unknown metallurgical variables, conflicting observations by different investigators and the distinction which has to be made between the diode surface and bulk changes. [37, 38]

The radiative efficiency in the spontaneous and lasing mode will be strongly affected by a decrease in η_i, while in the lasing mode, $\bar{\alpha}$ is a second important parameter. Because of the isotropic nature of the radiation in the incoherent mode of operation, the absorption coefficient generally differs from $\bar{\alpha}$ in that it represents the losses averaged over the bulk of the device ; i.e., in regions well

removed from the p-n junction.

If the density of nonradiative recombination centers is low, the internal quantum efficiency value relevant to LED operation (i.e., at low current densities) may differ from that relevant to laser diode operation where the injected carrier density at threshold is about 10^{18} cm^{-3} at 300 K. In addition, the radiative lifetime decreases with increasing current above lasing threshold further reducing the effects of a low density of nonradiative recombination centers.

Experimental evidence [39] that major changes in the *bulk* of the diode can occur during high current density operation ($J > 1000$ A/cm^2) was originally obtained in homojunction GaAs laser diodes prepared by liquid-phase epitaxy. The degraded laser output was shown to result from an internal efficiency decrease and increased absorption. Later results obtained using heterojunction diodes [40] have confirmed the fact that changes within the recombination region, rather than simple increase in the excess current, are mainly responsible for degradation. Detailed results obtained by various investigators differ, but the following observations on degradation appear to be generally valid.

1) The threshold current density of laser diodes increases while the differential quantum efficiency frequently (but not always) decreases.

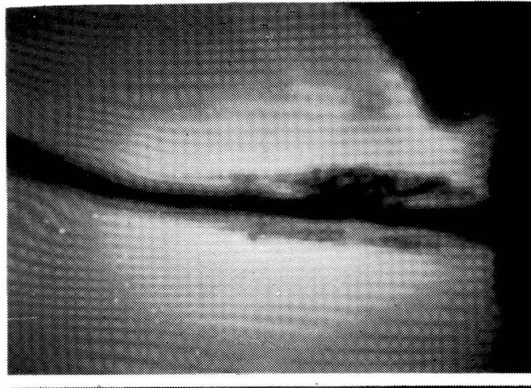
2) Heterojunction diodes operated at high current densities degrade similarly whether lasing or not, [40] (as long as facet damage does not occur) because the degradation rate is a function of the operating current density.

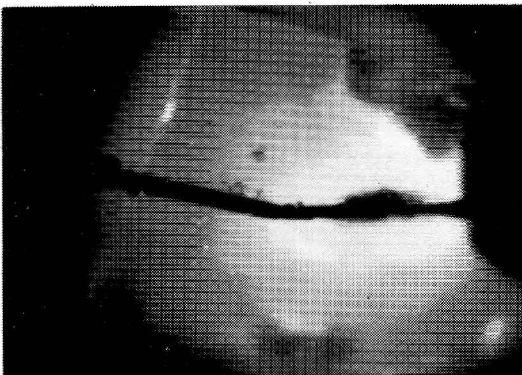
3) The minority carrier lifetime decreases with the spontaneous efficiency. [41] Assuming that the external spontaneous efficiency change follows that of η_i, the change of τ with η_i is consistent with Eq. 6 and can be attributed to a reduction in τ_{nr} (with τ_r remaining constant) because of the gradual formation of nonradiative recombination centers in the vicinity of the p-n junction.

Although τ and J_{th} are also found to be related, [42] an exact correlation between the reduction in τ and the increase in J_{th} is not always possible because any defects created may also increase $\bar{\alpha}$. [38,43]

4) Degradation does not require the presence of a p-n junction, but has been observed in heterotransition material (all n-type) *optically* pumped at a level either above or below lasing threshold. [44] No evidence of damage was observed in the transparent (AlGa)As window region of the structure indicating that electron-hole recombination was necessary for degradation to occur.

5) There are successive degradation stages. In the early stages, when the density of newly formed nonradiative recombination centers is low, the *spontaneous* emission efficiency decreases without any obvious changes in the emission pattern and with only minor effects on the *lasing* properties. As degradation proceeds, dark lines (commonly ⟨100⟩ oriented) may appear (when emission is viewed through the surface of the diode) as shown in Fig. 14(c), which are not evident in the early degradation stages, Fig. 14 (a,b). The dark lines constitute regions of concentrated nonradiative centers and





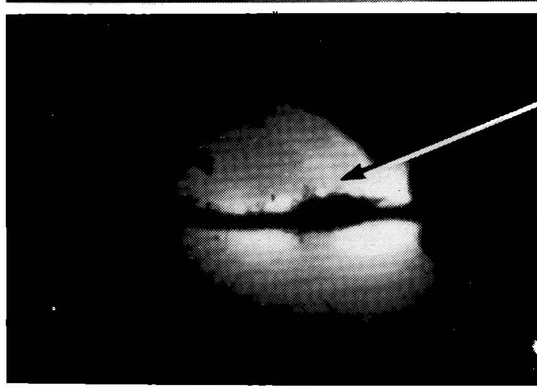
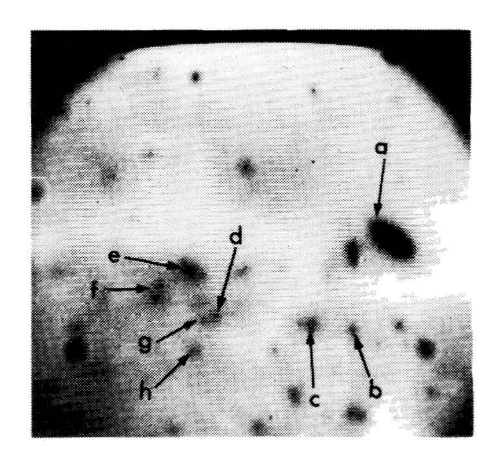


Figure 14 :
Effect of continuing operation on the surface emission from a double heterojunction diode. (a) Microphotograph of the infrared emission in the vicinity of the ohmic contact showing a clear emission pattern. (b) The diode output is reduced to I/5 of the initial value P_0 *, but no evidence of visible damage is seen. (The spots are on the surface of the diode). (c) The diode output is reduced to I/20 its initial value and dark lines are seen in the region of initial highest emission density. (Ref 38).*

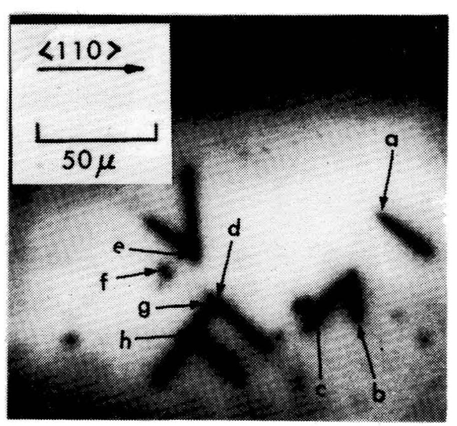
are believed responsible for the spotty near-field emission pattern of degraded lasers [39] and the *rapid* fall-off in the output of CW laser diodes at room temperature. [45]

These dark areas of concentrated nonradiative centers can obviously have several origins, including precipitation at existing dislocations. In the specific structures so far studied by transmission electron microscopy, evidence was found that the dark regions were newly generated dislocation networks formed by non-conservative climb from presumably starting at initially present dislocation sources. [46]

Studies have been made to identify the sources for these dark lines. In one case, at least, they have been found to originate in regions of the crystal which had originally low radiative efficiency as measured by photoluminescence, which could have propagated from the substrate into the epitaxial layer. [47] In fact, by using intense external laser photoexcitation, it has been possible to show the rapid growth of these dark regions as shown in Fig. 15.



(a)



(b)

Figure 15 :
Photomicrographs showing the growth of non-radiative " dark lines" as evidenced by photoluminescence before (a) and (b) after high intensity photoexcitation. (Ref. 47).

A number of important empirical observations have been made which have contributed to obtaining greatly improved devices. The degradation rate depends on many metallurgical factors for a given current density. Contaminants such as Cu are definitely detrimental in GaAs. Experiments with Cu in GaAs:Zn diffused homojunctions [48] yielded an activation energy for the degradation of 0.45± 0.1 eV with indications that the efficiency reduction was due to the internal quantum efficiency decrease rather than a simple increase in the

excess current. The observed activation energy for degradation being close to the activation energy for Cu diffusion in GaAs (0.5 eV), suggests that the movement of Cu is responsible for the degradation. However, Cu can be effectively removed by various bleaching techniques. For this reason it is unlikely Cu is a significant factor in controlling the operating life of devices made by liquid-phase epitaxy where the material is in contact with Ga, an excellent "getter" for Cu. However, other contaminants, so far unidentified, may be relevant.

Dislocations initially present in the crystal are detrimental, [49] as are dislocations which can easily be introduced in the material in the process of handling the material. There are also differences between acceptor dopants. [49] Care is needed in the assembly of diodes to prevent stresses being set up when "hard" solders are used to bond the diodes to the heat sink. In the stressed regions, dislocations can be generated in the course of diode operation because of the thermal gradients present in the material during high current density operation.

An important observation is that diodes in which the recombination region consists of (AlGa)As are less susceptible to gradual degradation than similarly grown devices with GaAs in the recombination region. [50] This is illustrated in Fig. 16 which shows the half-life of diodes operated at 1000 A/cm^2 (spontaneous emission) as a function of the emission wavelength peak (i.e, increasing Al content in the (AlGa)As recombination region). A variety of homojunction and heterojunction devices were tested and the general trend of improved life with increasing Al content is evident.

While a theoretical understanding of the fundamental phenomena involved in degradation still awaits further research, a number of important conclusions can be made which support the original hypothesis that nonradiative centers are formed, and that these could include vacancies. The fact that the nonconservative climb of dislocations has been observed suggests the presence of a high density of vacancies. Furthermore, the enhanced diffusion of contaminants such as Cu similarly indicates that vacancies are in some manner formed or move into the recombination region of the diode as a result of electron-hole recombination, although the mechanism for their formation is still in doubt. [38]

In view of the limited statistical data available so far concerning the operating life of LEDs and CW laser diodes, it is evidently not possible to conclusively state operating lifetimes useful in making systems decisions. It is, however, clear that the feasibility of fabricating LEDs operating at 1000 A/cm^2 with lifetimes in excess of 10.000 hours has been established with laboratory devices (see for example, Fig 16). With regard to laser diodes, a few samples in several laboratories have reached 5000 hours of operating life (although not with constant output) with emission of a few milliwatts of power, but much work remains to obtain reproducible devices. In addition to efforts directed toward improving the metallurgical perfection of the material grown and the use of (AlGa)As in the recombination region, some work has also been reported [51] in which the strain introduced by the difference in the

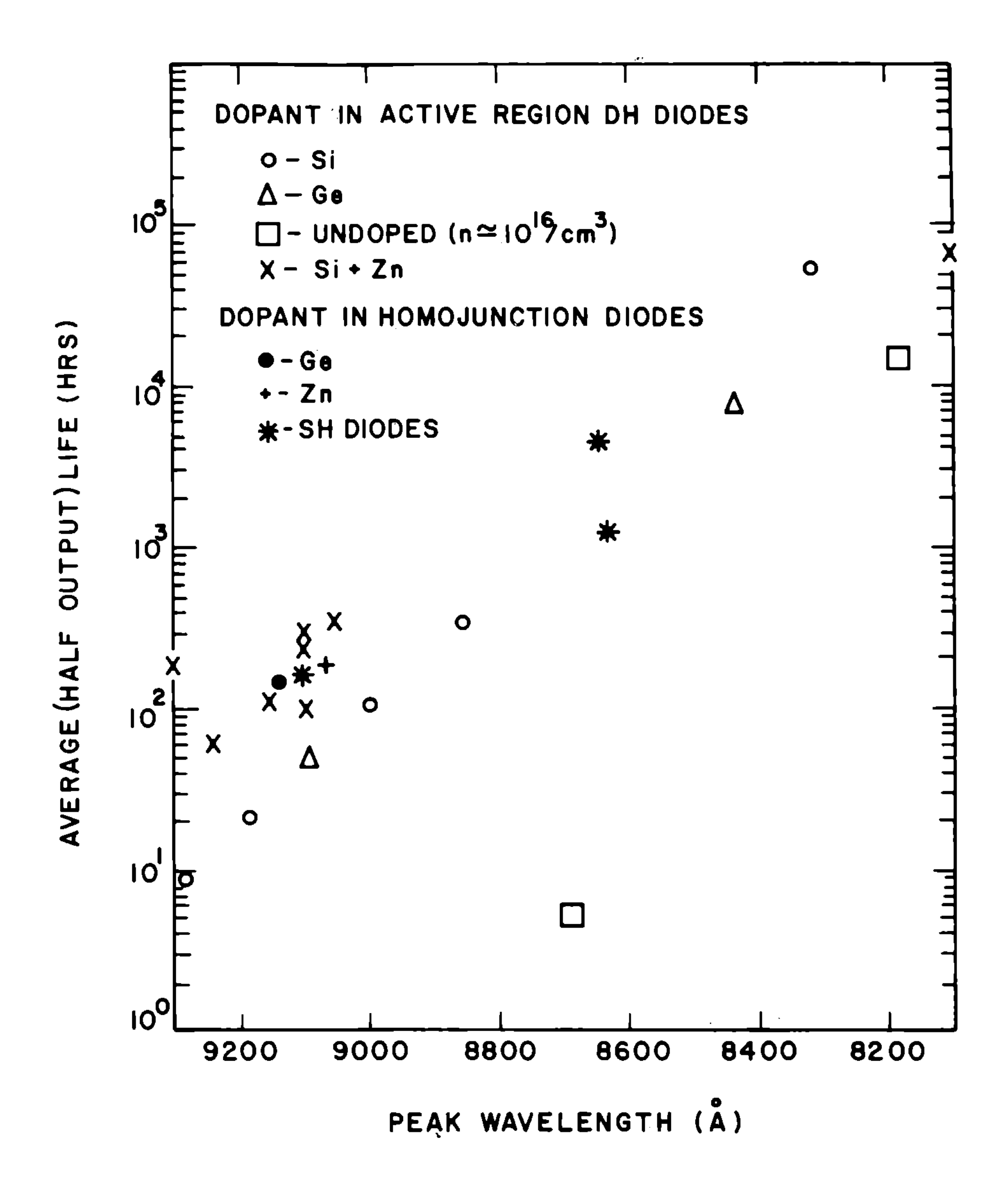


Figure 16 :
Average half-life of diodes operating in the incoherent mode at 1000 A/cm as a function of their emission wavelength. For each dopant type in the recombination region the shift to shorter wavelength represents additions of Al to the recombination region. Both single, double heterojunction and homojunction diodes were tested as indicated. (Refs. 30,50).

thermal coefficients of expansion of GaAs and (AlGa)As is reduced by the addition of P to the (AlGa)As. This process leads to the formation of misfit dislocations to accommodate the lattice parameter difference at the growth temperature.

■ REFERENCES

[1] H. Kressel and H. Nelson, RCA Review 30, 106 (1969)

[2] I. Hayashi, M.B. Panish, and P.W. Foy , I.E.E.E. J. Quantum Electron 5, 211 (1969).

[3] Zh. I. Alferov, V.M. Andreev, E.L. Portnoi, and M.K. Trukhan, 3, 1328 (1969). Trans. : Sov. Phys. Semiconductors 3, 1107 (1970).

[4] For data concerning Corning fibers, see D.B. Keck, R.D. Maurer, and P.C. Schultz, Appl. Phys. Letters 22 , 307 (1973). At 8000 Å the attenuation is 6 db/km.

[5] A. Yariv, *Quantum Electronics,* John Wiley and Sons, Inc., New York, 1967 ; and J.I. Pankove, *Optical Processes in Semiconductors,* Prentice-Hall, Englewood Cliffs, N.J., 1971.

[6] H. Kressel, "Semiconductor Lasers" in *Lasers*, Vol 3, A.K. Levine and A.J. DeMaria, eds., (Marcel Dekker, N.Y., 1971) ; *Laser Handbook,* F.T. Arecchi and E.O. Schulz-DuBois, Eds., (North Holland Pub. Co., Amsterdam, 1972).

[7] P.G. Eliseev, Sov. J. Quantum Electron. 2 , 505 (1973).

[8] J.K. Butler, J. Appl. Phys. 42, 4447 (1971).

[9] H. Kressel, H.F. Lockwood, and J.K. Butler, J. Appl. Phys. 44, 4095 (1973).

[10] H.F. Lockwood, H. Kressel, H.S. Sommers, Jr., and F.Z. Hawrylo, Appl. Phys. Letters 17, 499 (1970) ; H. Kressel, H.F. Lockwood and F.Z. Hawrylo, J. Appl. Phys. 42 , 561 (1972).

[11] G.H.B. Thomson and P.A. Kirkby, I.E.E.E. J. Quantum Electron. 9 , 311 (1973).

[12] F. Stern, I.E.E.E. J. Quantum Electron. 9, 290 (1973).

[13] P.T. Landsberg and M.J. Adams, in *The Physics and Technology of Semiconductor Light Emitters and Detectors,* A. Frova, ed., (North-Holland Publishing Co,. Amsterdam 1973) P.3.

[14] H. Kressel and H.F. Lockwood, Appl. Phys. Letters 20, 175 (1972).

[15] G.H.B. Thomson and P.A. Kirkby, Electronics Letters 9, 295 (1973)

[16] J.K. Butler, H.S. Sommers, Jr., and H. Kressel, Appl. Phys. Letters 17, 403 (1970).

[17] H. Kressel, J.K. Butler, F.Z. Hawrylo, H.F. Lockwood, and M. Ettenberg, R.C.A. Review 32 , 393 (1971).

[18] H. Kressel and H. Nelson, "Liquid Phase Epitaxy" in *Physics of Thin Films,* Vol 7, G. Hass, M.H. Francombe, and R.W. Hoffman, eds., (Academic Press, N.Y., 1973).

[19] H.F. Lockwood and H. Kressel, J. Crystal Growth, 1974 (in press)

[20] H. Nelson, U.S. Patent No. 3, 565, 702 (1971) M.B. Panish, I. Hayashi, and S. Sumski, Appl. Phys. Letters 16, 320 (1970)

[21] H.F. Lockwood and M. Ettenberg, J. Crystal Growth 15 , 81 (1972).

[22] J.C. Dyment and L.A. D'Asaro, Appl. Phys. Letters 11 , 292 (1967).

[23] T. Tsukada, H. Nakashima, J. Umeda, and D. Nakada, Appl. Phys. Letters 20 , 344 (1972).

[24] J.C. Dyment, L.A. D'Asaro, J.C. North, B.I. Miller, and J.E. Ripper, Proc. I.E.E.E. (Letters), 60 , 726 (1972).

[25] H. Yonezu, I. Sakuma, K. Kobayashi, T. Kamejima, M. Urno, and Y. Nannichi, Jap. J. Appl. Phys. 12 , 1585 (1973).

[26] M. Cross and M.J. Adams, Solid-State Electron. 15 , 919 (1972).

[27] B.W. Hakki, J. Appl. Phys. 44 , 5021 (1973).

[28] M. Chown, A.R. Goodwin, D.F. Lovelace, G.B.H. Thompson, and P.R. Selway, Electronics Letters 9, (25 January) 1973.

[29] H. Kressel and M. Ettenberg, Appl. Phys. Letters 23 , 511 (1973).

[30] M. Ettenberg, H.F. Lockwood, J.P. Wittke, and H. Kressel, *Technical Digest,* International Electron Devices Meeting, Washington, D.C., 1973 (I.E.E.E. Catalog No. 73CHO781-SED).

[31] J. Conti and M.J.O. Strutt, I.E.E.E. J. Quantum Electron. QE-8 , 815 (1972).

[32] T.P. Lee and C.A. Burrus, I.E.E.E. J. Quantum Electron. QE-8, 370 (1972).

[33] C.A. Burrus and R.W. Dawson, Appl. Phys. Letters 17, 97 (1970).

[34] M. Ettenberg, K. Hudson, and H.F. Lockwood, I.E.E.E. J. Quantum Electron. QE-9, 987 (1973).

[35] J.P. Wittke , private communication.

[36] I. Ladany and H. Kressel, to be published.

[37] P.G. Elisev, *The Physics and Technology of Semiconductor Light Emitters and Detectors,* A. Frova , ed., (North-Holland Publishing Co. , Amsterdam 1973), p. 338.

[38] For recent review see H. Kressel and H.F. Lockwood. J. de Physique, 1974 (in press).

[39] H. Kressel and N.E. Byer, Proc. I.E.E.E. 38, 25 (1969).

[40] D.H. Newman, S. Ritchie, and S. O'Hara, I.E.E.E. J. Quantum Electron. QE-8, 379 (1972).

[41] E.S. Yang, J. Appl. Phys. 42, 5635 (1971).

[42] N. Chinone, R.Ito, and O. Nakada, I.E.E.E. J. Quantum Electron. QE-10, 81 (1974).

[43] B.W. Hakki and T.L. Paoli, J. Appl. Phys. 44, 4113 (1973).

[44] W.D. Johnston and B.I. Miller, Appl. Phys. Letters 23 , 192 (1973).

[45] B.C. Deloach, B.W. Hakki, R.L. Hartman, and L.A. D'Asaro, Proc. I.E.E.E. 61, 1042 (1973).

[46] P. Petroff and R.L. Hartman, Appl. Phys. Letters 23 , 469 (1973).

[47] R. Ito, H. Nakashima, and O. Nakada , Jap. J. Appl. Phys., 1974 (in press).

[48] A. Bahraman and W.G. Oldham, J. Appl. Phys. 43 , 2383 (1972).

[49] H. Kressel, N.E. Byer, H.F. Lockwood, F.Z. Hawrylo, H. Nelson, M.S. Abrahams, and S.H. McFarlane, Met. Trans. 1, 635 (1970).

[50] M. Ettenberg, H. Kressel, and H.F. Lockwood, Appl. Phys. Letters, July 1974.

[51] J.C. Dyment, F.R. Nash , C.J. Hwang, G.A. Rozgonyi, R.L. Hartman, H.M. Marcos, and S.E. Haszko, Appl. Phys. Letters 24 , 481 (1974).

PROPRIETES DES DIODES ELECTROLUMINESCENTES

J. LEBAILLY
RTC La Radiotechnique-Compelec
14000 Caen France

■ INTRODUCTION

Nous nous proposons de décrire les propriétés des sources peu intenses que constituent les diodes électroluminescentes utilisées sous densité de courant modérée , c'est à dire inférieure à 1000 A/cm^2, dans une perspective d'application en photonique. Par opposition, nous considérons comme "sources intenses" les diodes électroluminescentes fonctionnant sous densité de courant supérieure à 1000 A/cm^2 (en particulier les diodes de structure dite de Burrus |1|) et les lasers à semiconducteurs.

Afin de restreindre le sujet, nous ne considérerons que les dispositifs envisageables dans les deux cas suivants:-systèmes monolithiques d'optique intégrée, réunissant par exemple source, modulateur ou commutateur, récepteur;-sources destinées aux communications par fibre optique multimode. Cette dernière application limite les sources à considérer à celles dont le spectre d'émission se situe dans l'une des fenêtres de transparence des fibres, 800 à 850 nm et 1000 à 1100 nm. Cette dernière bande est parfaitement compatible avec les lasers YAG dopés au Nd ou ultraphosphate de néodyme, ainsi qu'avec un certain nombre de sources à semiconducteurs d'introduction plus récente (au GaInAs en particulier). La bande 800-850 nm constitue le domaine de prédilection des sources à semiconducteurs: diodes électrolumiscentes et lasers au GaAlAs.

■ II. VIEILLISSEMENT EN FONCTIONNEMENT

La dégradation irréversible, en fonctionnement, des diodes électroluminescentes et plus encore des lasers constitue une limitation fondamentale non encore résolue. S'il existe des lasers dont la durée de fonctionnement dépasse non plus quelques heures mais quelques milliers d'heures, il n'en demeure pas moins vrai qu'une telle durée de vie est incompatible avec des applications

telles que la téléphonie. Nous verrons que les diodes électroluminescentes présentent une durée de fonctionnement élevée à condition qu'elles soient utilisées sous faible densité de courant. Mais dans ces conditions le niveau de signal émis apparait souvent "tout juste utilisable".

Différents mécanismes de dégradation sont maintenant connus, et il apparait que ceux-ci sont les mêmes , à des degrés différents, pour les lasers et les diodes électroluminescentes, à part quelques effets spécifiques aux premiers tels que la destruction des faces réfléchissantes du laser.

Pour l'utilisateur de diodes électroluminescentes, le vieillissement du dispositif peut se traduire de deux façons sur la caractéristique puissance lumineuse émise-courant injecté représentée Figure 1 en unités log-log. On peut observer séparément ou simultanément soit une décroissance de l'efficacité quantique de luminescence (glissement vers le bas de la tangente à 45°), soit un accroissement de composante(s) de courant non radiative(s) (déformation de la caractéristique à bas niveau). Dans les diodes au GaAs l'effet prépondérant se trouve être le second. Dans ces conditions, le procédé habituel est incorrect, qui consiste à observer la valeur du signal émis sous la valeur nominale du courant de fonctionnement et à extrapoler l'évolution observée. Il est au contraire indispensable d'analyser et d'extrapoler séparément les diverses composantes de dégradation |2|.

A. Evolutions rapides: contraintes mécaniques et thermiques

Le premier mécanisme de dégradation auquel se heurtent les fabricants de diodes électroluminescentes (ou de lasers) provient des contraintes mécaniques et thermiques développées dans le cristal au cours du fonctionnement. Considérons une diode électroluminescente travaillant sous densité de courant modérée, 100 A/cm^2. Si l'essentiel de la puissance est dissipée dans une couche d'épaisseur de l'ordre de 5 µ (étendue de la région dans laquelle ont lieu les recombinaisons et l'absorption des photons; cas du GaAs ou du GaAlAs), la valeur de la puissance thermique dissipée par unité de volume de cette couche est de

$$(300\ A/cm^2 \times 1{,}4\ V)/(5.\ 10^{-4} cm) \sim 10^6\ W/cm^3.$$

Dans le cas d'un laser elle atteint 10^8 W/cm^3. Dans ces conditions le cristal peut être soumis à des gradients thermiques et à des contraintes mécaniques excessives, d'autant plus que les propriétés mécaniques et thermiques des composés III-V sont plus défavorables que celles du silicium. On observe alors une décroissance relativement rapide de la puissance lumineuse émise par la diode électroluminescente (en quelques heures par exemple) ou par le laser analysé en régime d'émission spontanée. Ceci s'accompagne de l'apparition de "lignes noires" à faible durée de vie de porteurs libres, orientées dans des directions cristallines correspondant aux plans de clivage préférentiels |3| . Il a été montré que ces effets prennent des proportions catastrophiques -lorsque les contraintes provoquées par un dépôt diélectrique ou par une forte concentration d'impuretés introduites par diffusion sont excessives |3|-lorsque le cristal

est fixé sur un dissipateur de coefficient de dilatation différent et à l'aide d'une soudure dure |4|-lorsque le cristal est soumis à une pression uniaxiale lors de son fonctionnement |5| -ou lorsque

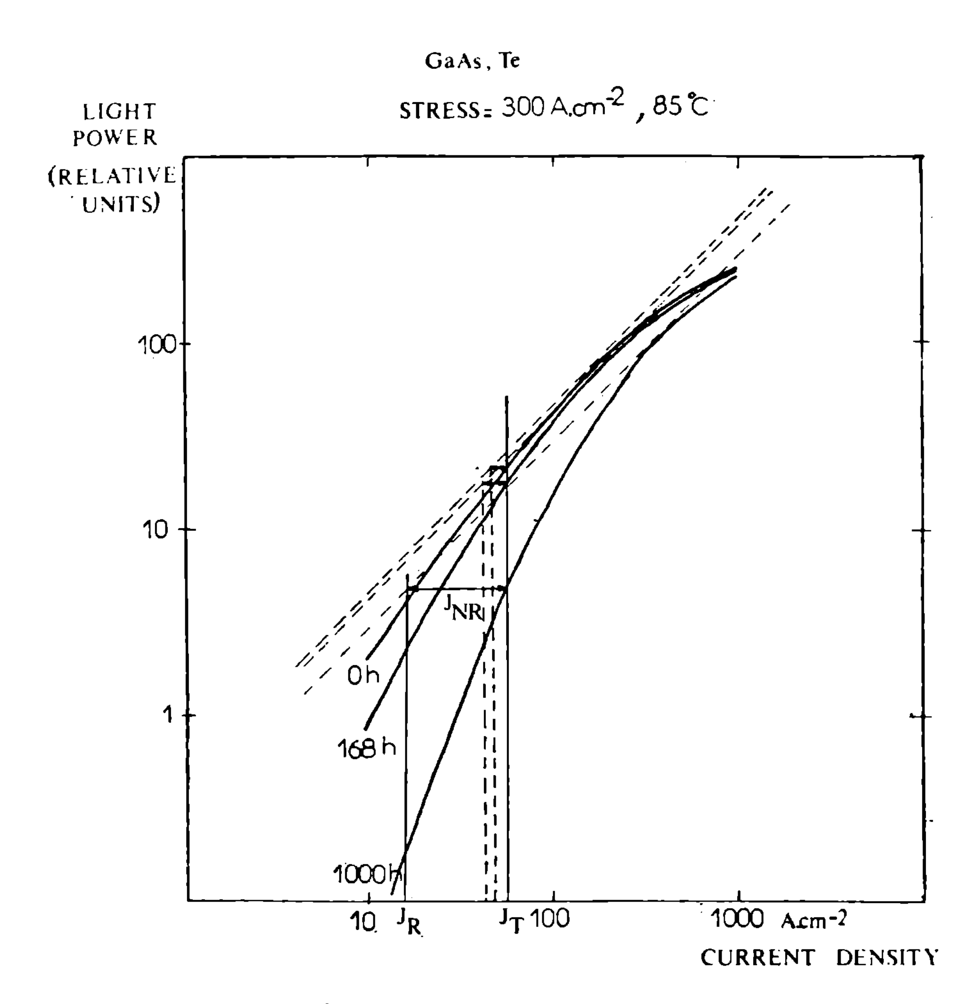


Figure 1:
Caractéristique puissance lumineuse émise-courant injecté d'une diode au GaAs dopée au Te, et son évolution en vieillissement. La position de la tangente à 45° définit l'efficacité quantique d'émission radiative. La nonlinéarité à bas niveau est attribuée à une composante de courant non radiative notée J_{nr}. La nonlinéarité à haut niveau est due à un effet thermique. Dans le cas présent, on constate que la dégradation se traduit par une décroissance de l'efficacité quantique et, surtout, par un accroissement de la composante de courant non radiative.

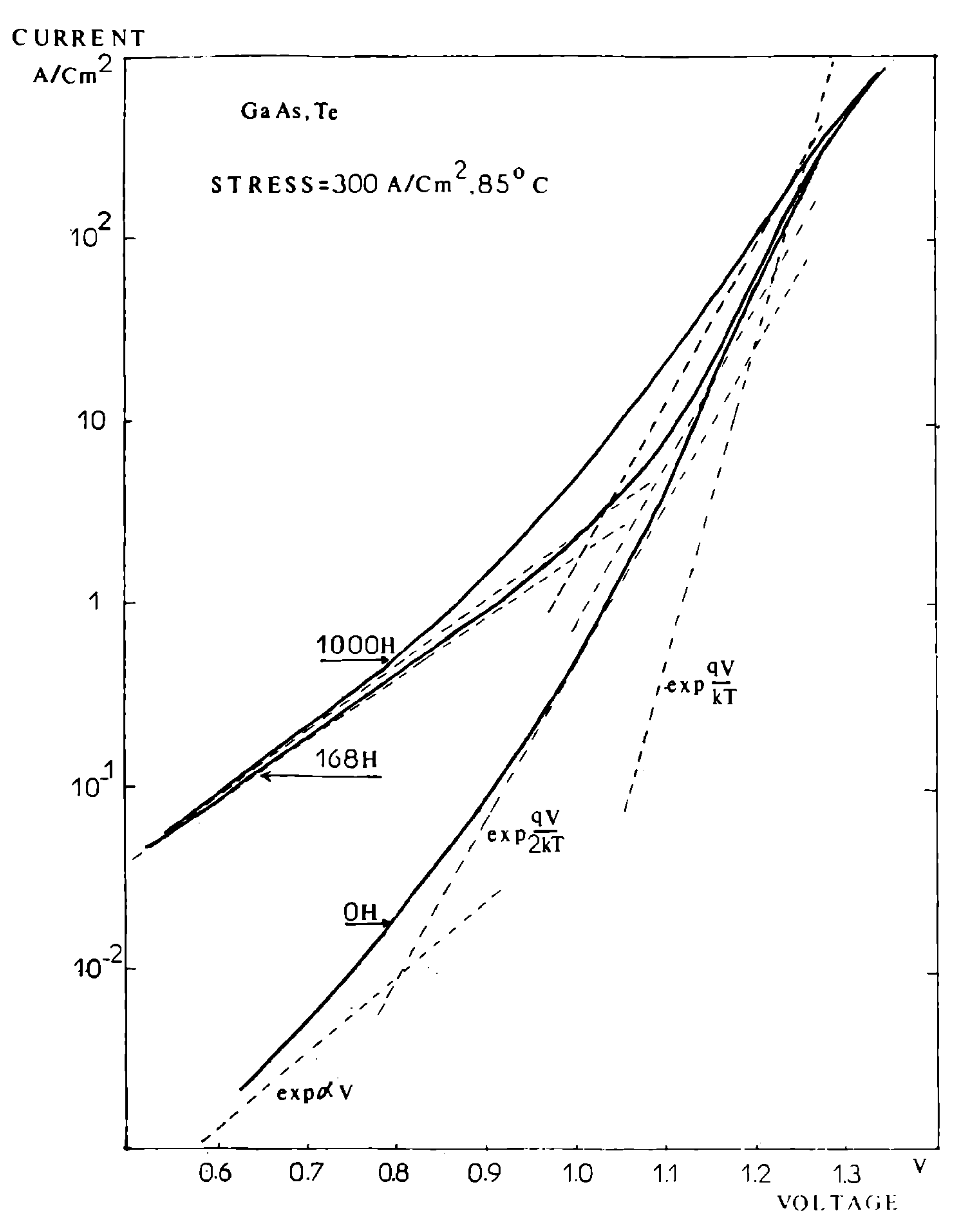


Figure 2:
Caractéristique courant-tension d'une diode au GaAs dopée au tellure, et son évolution en vieillissement. On constate que la composante de courant non radiative est dans ce cas la somme d'un terme exp(qV/2kT) qui croît sensiblement et d'un terme en exp(α V) qui croît considérablement au cours de la période initiale de vieillissement.

la densité de courant est répartie de façon non uniforme |6|. Les différentes observations ont permis d'aboutir à des structures de dispositifs améliorées sinon optimisées du point de vue des contraintes.

B. Evolutions lentes liées à la qualité du matériau

Après cette éventuelle décroissance rapide apparait une évolution lente de la puissance lumineuse émise. Celle-ci a fait l'objet d'études systématiques de notre part dans le cas du GaAs |2|. Il a été montré que:
-la dégradation se manifeste d'une part par la croissance de composantes

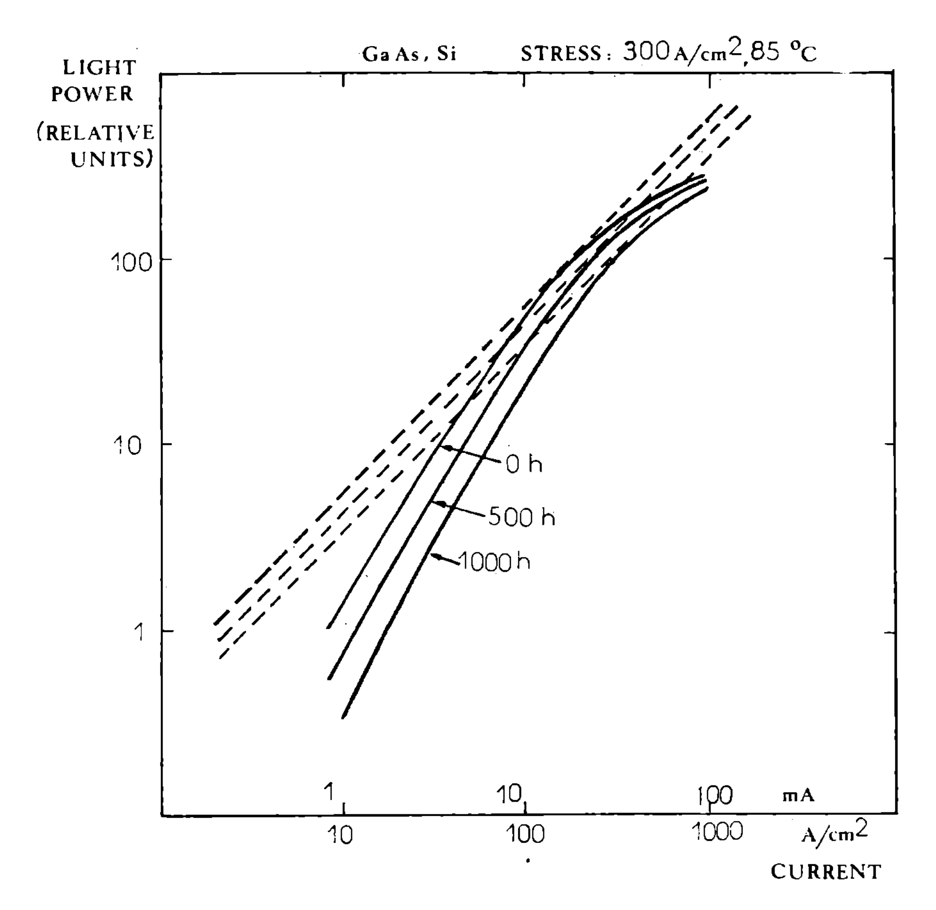


Figure 3:
Même type de caractéristique que pour la Figure 1, avec dans le cas présent un matériau GaAs dopé au silicium, et une dégradation se traduisant par une décroissance de l'efficacité quantique et un accroissement de la composante de courant non radiative, dans des proportions comparables.

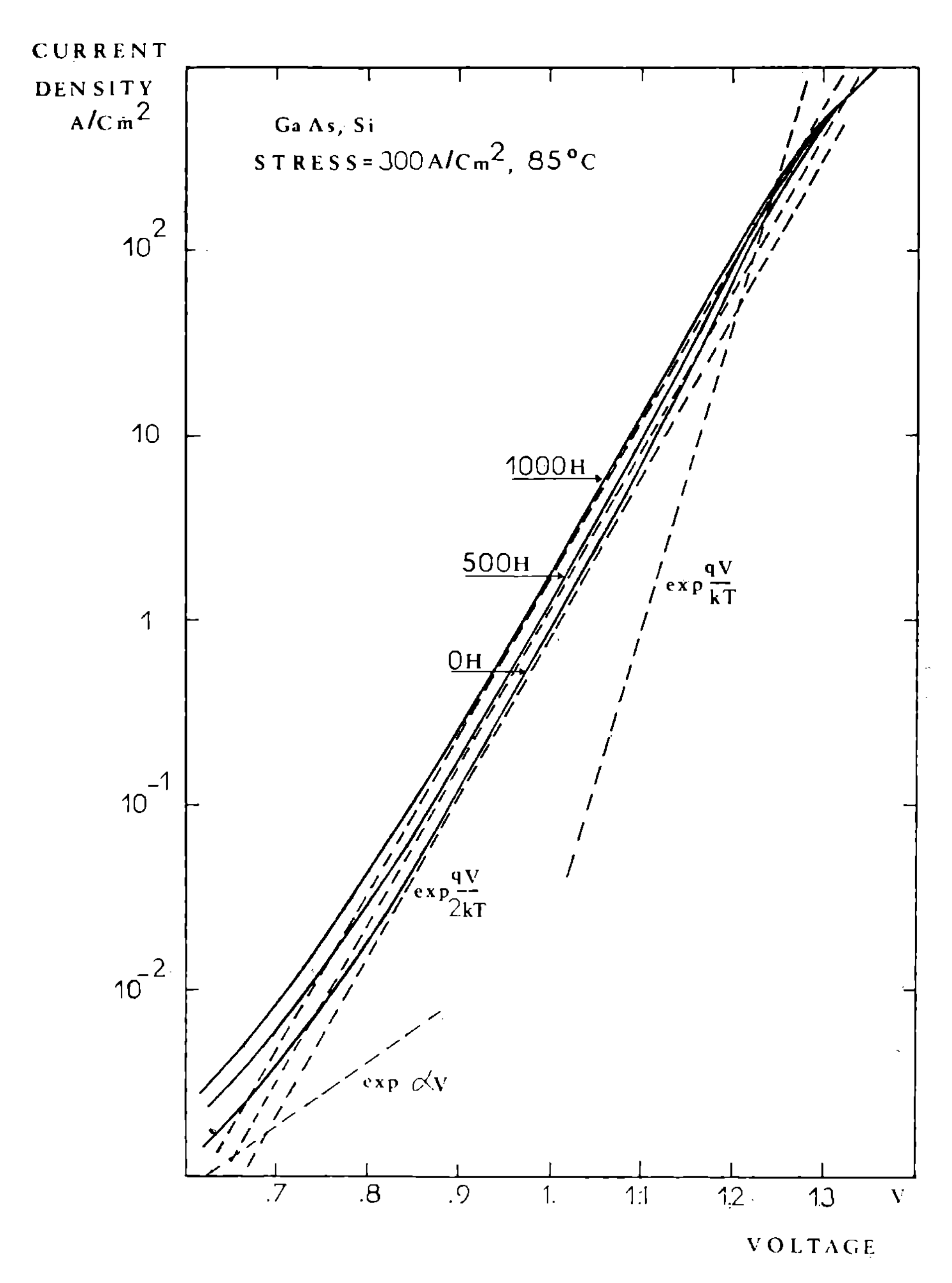


Figure 4:
Même type de caractéristique que pour la Figure 2, avec dans le cas présent un matériau GaAs dopé au silicium. La composante non radiative en qV/2kT est prépondérante dans ce cas.

de courant totalement non radiatives (courant de recombinaison indirecte dans la région de charge d'espace, courant tunnel par l'intermédiaire de centres profonds); d'autre part en proportions moindres

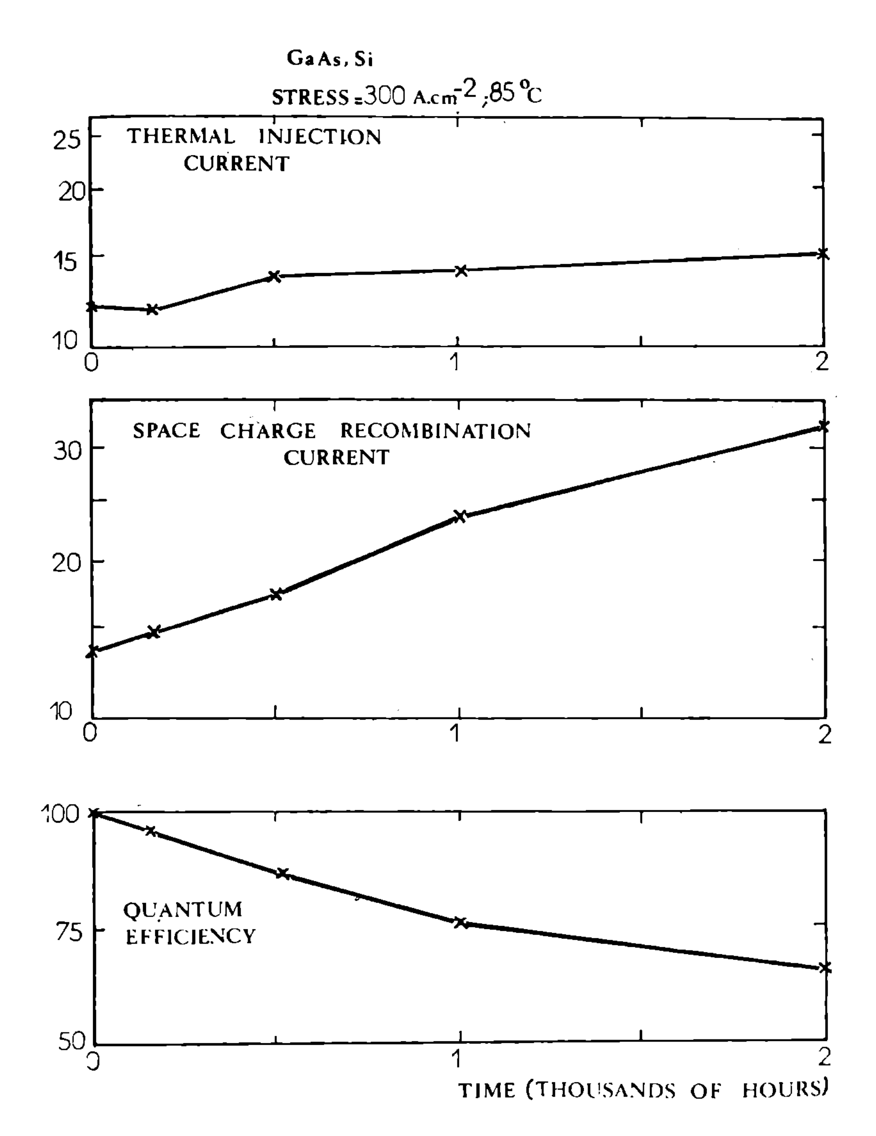


Figure 5
L'extrapolation des évolutions de chaque terme intervenant dans les caractéristiques du dispositif est possible et nécessaire si l'on veut procéder à une estimation correcte de la durée d'utilisation de la diode. Les évolutions représentées ici correspondent au cas des Figures 3 et 4.

par une décroissance de l'efficacité quantique de luminescence de la composante de courant d'injection thermique (composante radiative). Figures 1 à 4.
-les évolutions de chacun des paramètres sont extrapolables et permettent d'estimer de façon valable la durée de fonctionnement du dispositif. Figure 5.
-la nature du dopant (site cristallin sur lequel il se place en substitution, rayon atomique) intervient dans la façon dont se manifeste la dégradation (en particulier la nature des composantes de courant non radiatives).
-la qualité du matériau est le paramètre prépondérant; la densité de dislocations est un critère de qualité du point de vue des dégradations insuffisant cependant.
-les évolutions observées sont compatibles avec l'hypothèse d'une certaine probabilité de génération de centres de recombinaison non radiative lors des recombinaisons de porteurs |2|; cependant la faible valeur de l'énergie correspondante et le fait que la fragilité apparait très variable d'un lingot à un autre lingot ou d'une plaque épitaxiale à une autre plaque épitaxiale conduisent à supposer que l'on a affaire non pas à une génération de défauts de Frenkel mais plutôt à une dissociation ou à une constitution de complexes mettant en jeu des défauts présents initialement.

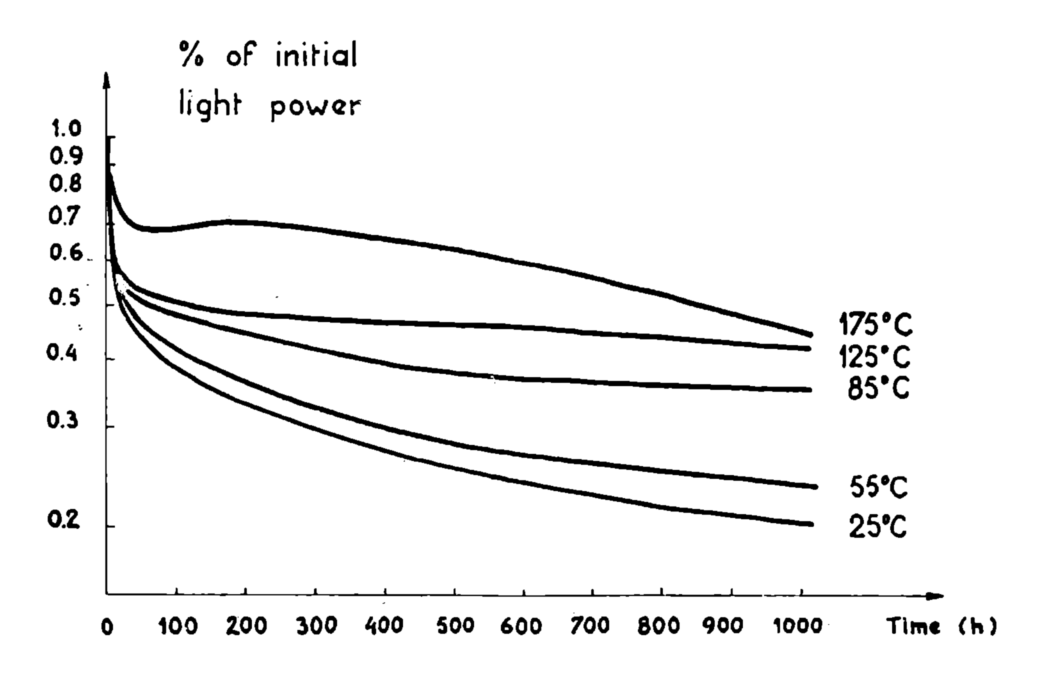


Figure 6:
Exemple de comportement "anormal" dans le cas d'un cristal initialement contaminé par des ions susceptibles de migrer d'autant plus aisément que la température est élevée.

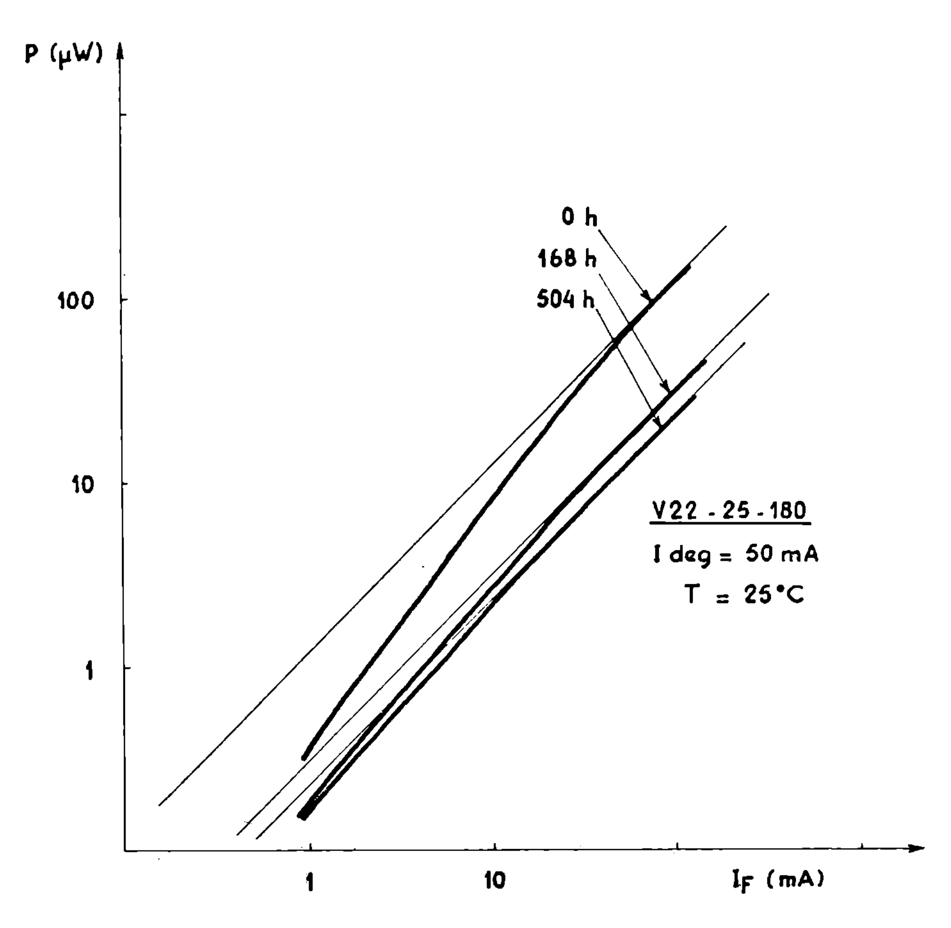


Figure 7:
L'importance de l'évolution de la caractéristique puissance lumineuse - courant confirme l'amplitude des dégradations observées dans le cas de la Figure 6.

Dans le cas d'une surface de jonction relativement grande ($\geq 10^4$ cm^2) et d'une structure non optimisée du point de vue des contraintes mécaniques, on atteint en général une puissance lumineuse émise moitié de la valeur de départ au bout de 10^4 heures de fonctionnement sous densité de courant de 100 A/cm^2 et 10^5 heures sous 10 A/cm^2. Il est probable que ces valeurs peuvent être nettement améliorées d'une part avec des structures de plus petite surface, d'autre part avec des matériaux de qualité accrue.

C. Evolutions anormales attribuées à des polluants

Le type décrit ci-dessous a été observé avec un lot particulier de plaques épitaxiales de $GaAs_{0,6}P_{0,4}$ dopées au sélénium et présentant des concentrations anormalement élevées d'impuretés ioniques

susceptibles de migrer dans le cristal. Il représente un cas extrême et permet de schématiser le comportement résultant de la migration d'impuretés indésirables. Une première anomalie apparait dans le comportement en fonction de la température, la dégradation étant plus importante à basse qu'à haute température (Figure 6). L'évolution de la caractéristique puissance lumineuse émise-courant injecté laisserait croire à une forte diminution de l'efficacité quantique de luminescence (Figure 7). En réalité il apparait que la composante de courant d'injection thermique croît notablement (Figure 8). Une analyse complémentaire fait apparaitre une chute de la durée de vie des trous libres minoritaires au voisinage de la jonction. L'interprétation retenue est que l'on a affaire à une forte réduction de la durée de vie des porteurs minoritaires côté n de la jonction p-n, qui affecte en premier lieu l'efficacité d'injection d'électrons dans la région de type p luminescente. Ce type de dégradation a également été mis en évidence avec le (GaP,ZnO) |6| . Nous avons observé qu'il est partiellement réversible et peut être réduit par un dépôt diélectrique au rôle de "getter".

La contamination doit donc être évitée non seulement lors de la croissance des matériaux mais aussi lors de l'élaboration des dispositifs, à l'aide d'une passivation qui évite le contact entre la région active du cristal et le milieu extérieur souvent constitué par un matériau plastique.

D. Conséquence: règles de conception

En conclusion, la structure d'une source doit respecter les conditions suivantes:

1) Les contraintes d'ordre mécanique et thermique au sein du cristal doivent être réduitesau minimum. Même pour les sources intenses, ce problème est en passe d'être résolu de façon acceptable par l'utilisation de soudures douces ou contacts mous, de géométrie optimisée des jonctions diffusées et épitaxiales, de composition des couches épitaxiales (solutions ternaires et quaternaires|7|) telle qu'il n'y ait pas de différence de maille cristalline à température ambiante.

2) La qualité du matériau et, en fonction de celle-ci, la densité de courant de fonctionnement doivent être compatibles avec la durée d'utilisation visée. Dans l'état actuel des techniques, le matériau le plus pur est obtenu par épitaxie en solution de gallium, et, dans des conditions d'extrapolation correcte, une durée d'utilisation supérieure à 10^4 heures n'est compatible qu'avec une densité de courant inférieure à 10^3 A/cm^2.

3) Les dispositifs doivent être passivés de telle sorte que la surface ne puisse être polluée par les agents extérieurs et que les courants de surface demeurent réduits au minimum. Un cas idéal, vers lequel on doit tendre, est une structure dans laquelle la jonction active n'atteint la surface en aucun point et la surface libre est totalement enrobée d'une couche diélectrique de protection.

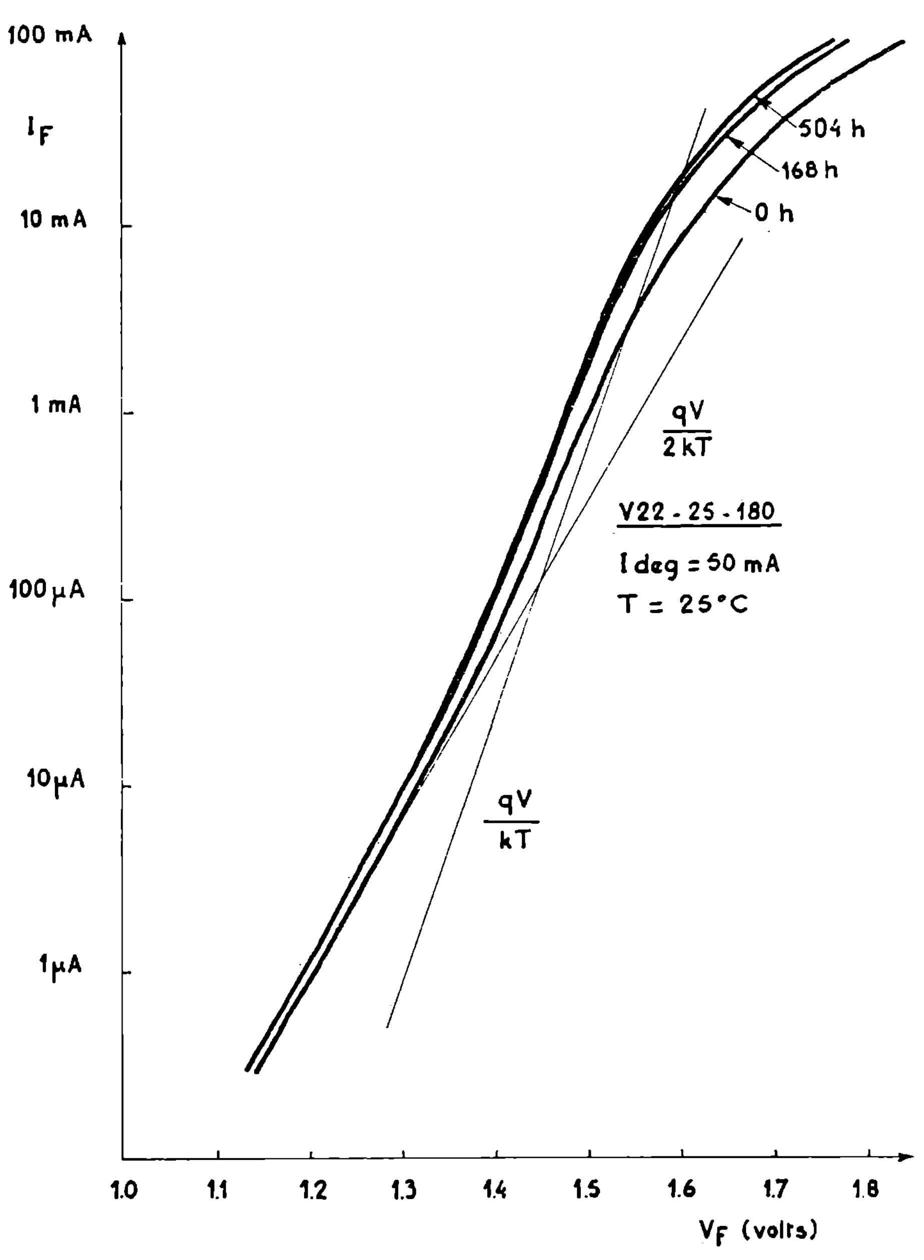
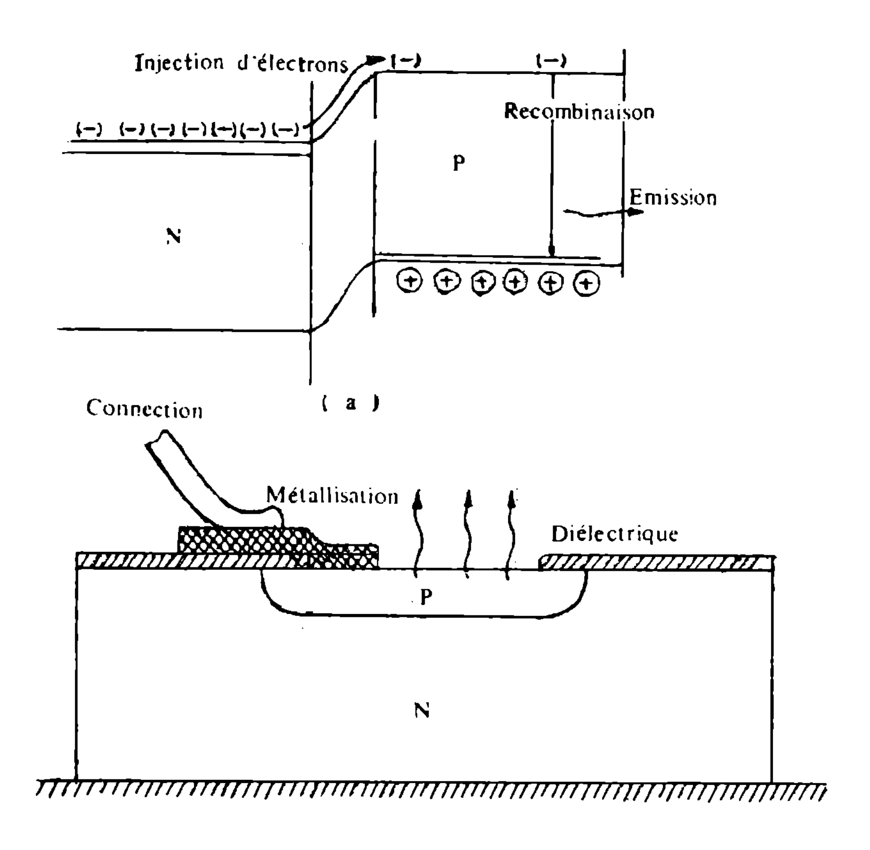


Figure 8:
L'analyse de la caractéristique courant-tension correspondant au cas des Figures 6 et 7 montre un accroissement du courant d'injection thermique dont la composante électronique est radiative. La réduction de la durée de vie des porteurs minoritaires côté n de la jonction est si importante qu'il apparait un courant important de trous injectés côté n.



(b)

Figure 9:
Structure classique d'une diode électroluminescente.

III. DIODE PLANE A CONFINEMENT DE PORTEURS

Le dispositif schématisé sur la Figure 10 représente une bonne approche de cette structure idéale. Dans celui-ci, les électrons libres, de faible masse effective, injectés du côté p de la jonction p-n sont confinés dans un volume dont on limite l'épaisseur à une valeur de l'ordre de 1 μ à l'aide d'une hétérojonction. Le matériau de faible largeur de bande interdite est du $Ga_{0,9}Al_{0,1}As$ de sorte que le spectre émis se situe dans la bande de transmission 800-850 nm des fibres optiques. Le matériau de grande largeur de bande interdite, permettant le confinement des électrons et constituant une fenêtre transparente est du $Ga_{0,7}Al_{0,3}As$. Dans ces conditions, aucun électron injecté ne peut atteindre la surface, le courant ne comportant pas de composante de surface se limite à la composante de courant d'injection thermique d'électrons et de

recombinaison radiative dans la région de confinement. Ceci conduit à une linéarité remarquable de la caractéristique de puissance lumineuse émise en fonction du courant injecté (Figure 11). Par ailleurs il y a lieu de souligner que le GaAlAs permet la formation d'une couche d'oxyde naturel, barrière efficace vis à vis des impuretés provenant du milieu ambiant en contact avec le cristal.

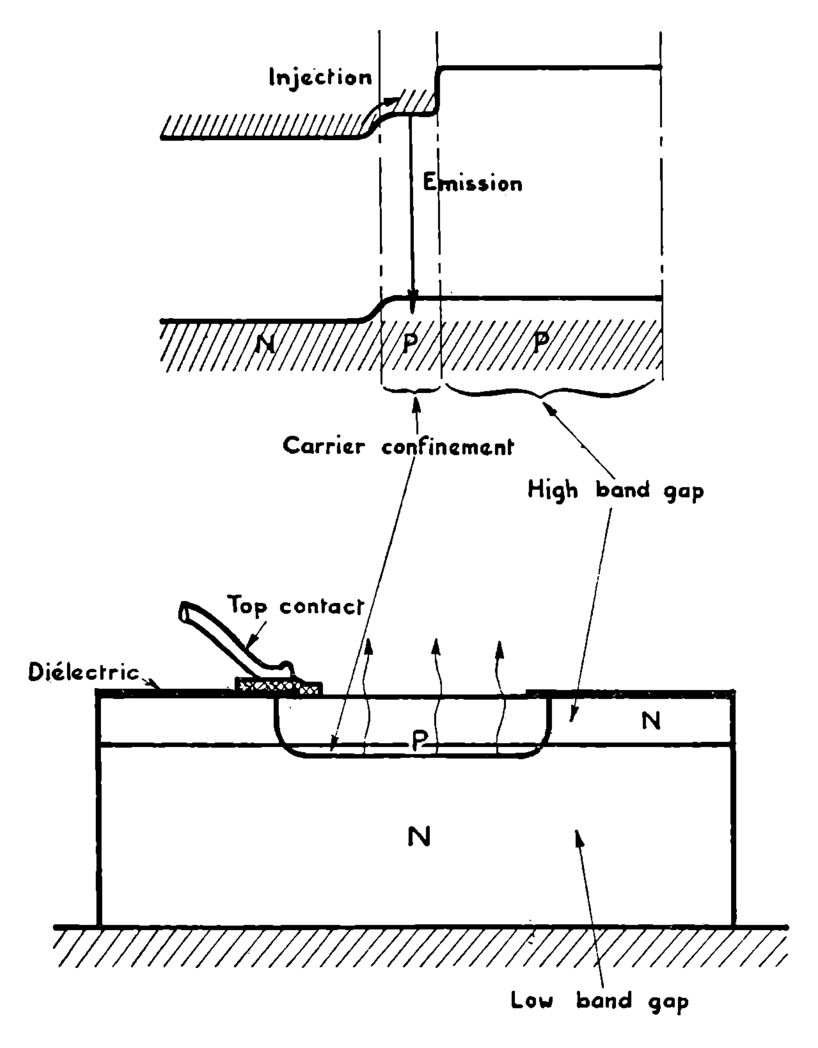


Figure 10:
Structure d'une diode à confinement de porteurs (à l'aide d'une simple hétérojonction) passivée.

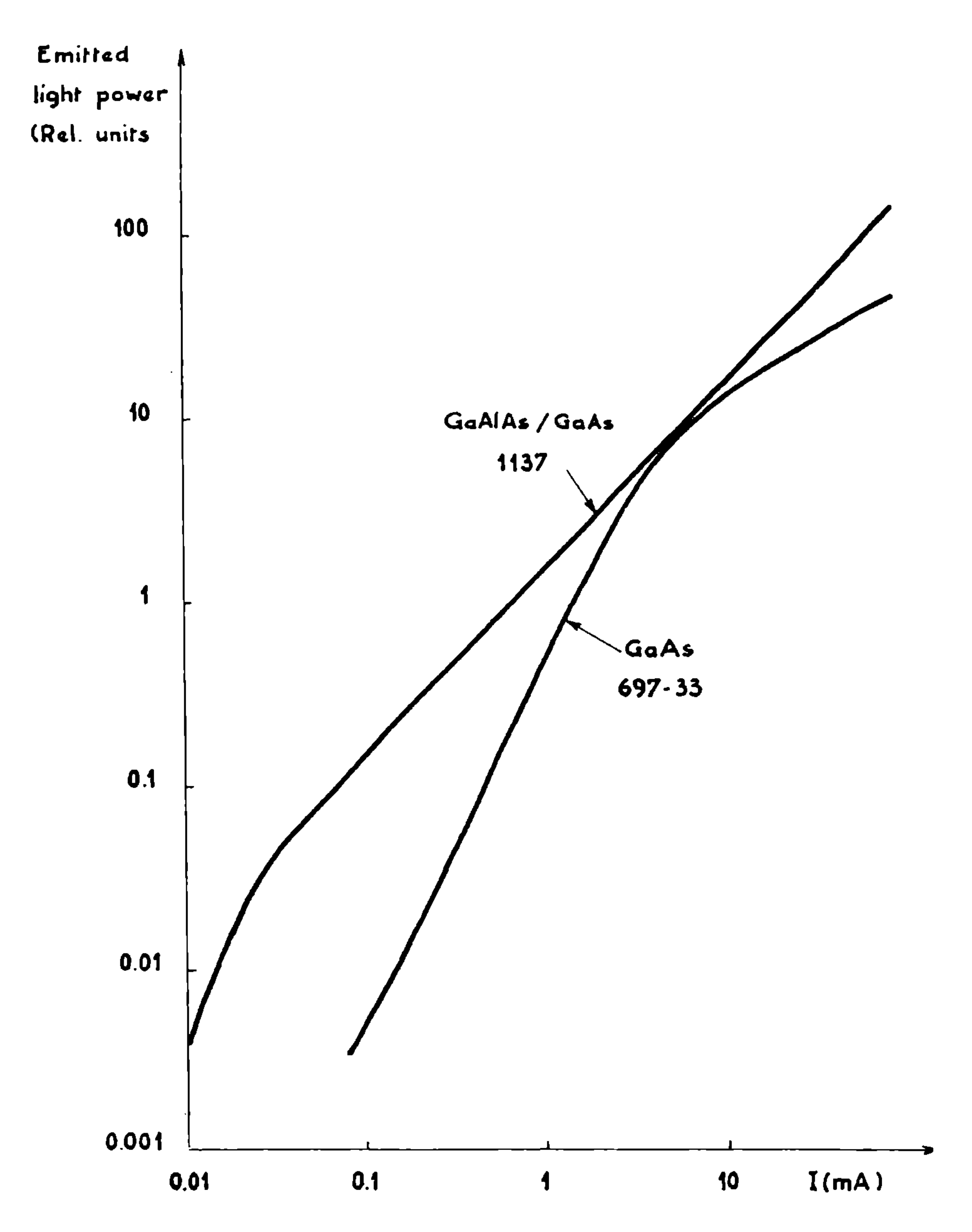


Figure 11:
Caractéristiques puissance lumineuse - courant comparées d'une diode au GaAs classique (notée 697-33) et d'une diode au GaAs à confinement de porteurs (notée 1137). La linéarité remarquable de cette dernière résulte de l'absence de courant de surface. (La hauteur de barrière de l'hétérojonction p-p empêche les électrons injectés côté p d'atteindre la surface).

A. <u>Propriétés comparées des différents types de sources électroluminescentes fonctionnant sous densité de courant modérée.</u>

Le tableau 1 résume les propriétés comparées des trois types de diodes suivants.

Les diodes au GaAs, élaborées par diffusion de Zn et utilisées couramment dans la construction de la plupart des photocoupleurs ou photoisolateurs. Le rendement quantique externe est limité à environ 1% malgré la valeur élevée du rendement quantique interne, environ 50%, par l'importance du mécanisme de réflexion totale des photons en cas d'indidence à un angle supérieur à 17° par rapport à la normale, et absorption après réflexion sur la surface (des photons ayant une énergie de l'ordre de la bande interdite du cristal) |8| .

Les diodes au GaAs dopées au Si amphotère sont élaborées par épitaxie en solution de telle façon que la jonction p-n est fortement dopée, compensée et graduelle. Dans ces conditions les recombinaisons ont lieu entre queue de bande de conduction et bande d'impuretés acceptrices, l'énergie des photons est inférieure à la largeur de la bande interdite, l'absorption est réduite dans

	λ Δλ (nm)	t (ns)	B(W/cm²/St) à 1000 A/cm²	P/1 (mW/A)
GaAs,Si 450 ×450 μ	930,40	200	9 13,3	28 (60%) 19 (40%)
GaAs,Zn 450 ×450 μ	880,25	20	4,6 4,6	13 (90%) 1,3 (10%)
GaA|As/GaA As, Zn	800 à 850,25	10	5 21	15 (90%) 1,5 (10%)

<u>Tableau 1</u>:
Comparaison des caractéristiques des trois types de diodes suivants: diode épitaxiale au GaAs, dopée au silicium amphotère (notée GaAs,Si); diode classique au GaAs, obtenue par diffusion de zinc (notée GaAs,Zn); diode à confinement de porteurs au GaAlAs (notée GaAlAs/GaAlAs). Le spectre émis est caractérisé par la longueur d'onde λ correspondant au maximum du spectre et sa largeur à mi-hauteur Δλ ; t est le temps de réponse du signal lumineux. Les valeurs de luminance B et de puissance lumineuse émise P correspondant aux émissions par la face supérieure et par le côté sont comparées.

un facteur de l'ordre de 5 pour un cristal en forme de parallélépipède rectangle.

Ces deux cas ne sont pas directement applicables aux télécommunications par fibres optiques en raison de la valeur de la longueur d'onde émise. Cependant le premier est donné car il représente la situation d'un matériau à structure de bande directe avec recombinaison radiative bande à bande. Les ordres de grandeurs cités pour le GaAs demeurent valables pour le $Ga_{1-x}Al_xAs$ de composition x inférieure à 0,2.

Le second cas, GaAs,Si, a été choisi non parce que de type de diode est largement utilisé en particulier pour la construction de lecteurs de bandes, cartes, mémoires mortes, etc .. mais pour la raison suivante. L'utilisation de recombinaisons bande à bande est incompatible avec l'intégration monolithique en optique intégrée en raison des valeurs élevées des coefficients d'absorption correspondant ($\gtrsim 10^3$ cm^{-1}). Par contre l'utilisation des transitions mettant en jeu les queues de bandes et bandes d'impuretés de substitution aux fortes concentrations peut constituer un moyen de réduire à un niveau acceptable l'absorption des photons émis. Ceci est illustré de façon simple par le rapport du nombre de photons émis par la face supérieure du cristal au nombre de photons émis par l'ensemble des quatre faces latérales: 90/10 et 60/40 respectivement.

B. Applications aux télécommunications par fibre optique multimode.

La diode à confinement de porteurs de structure plane constitue la "source peu intense" la mieux adaptée au couplage à une fibre multimode. Par rapport à une homojonction p-n, le confinement de porteurs introduit par l'hétérojonction se traduit par une densité d'électrons supérieure à la valeur moyenne qui résulte de l'injection thermique dans une homojonction. Il en résulte des propriétés semblables à celles que l'on observe à forte excitation dans une jonction p-n: à efficacité quantique égale, le temps de réponse est réduit dans un rapport voisin de 2 ou, à temps de réponse égal, l'efficacité d'électroluminescence est environ le double de celle d'une simple jonction p-n. Dans ces conditions, temps de montée et de descente des impulsions lumineuses $\sim$ 10 ns, efficacité quantique externe $\sim$1% (c'est à dire puissance lumineuse émise $\sim$15 mW/A), considérons une diode dont la surface émissive corresponde au diamètre d'une fibre multimode de diamètre 50 microns, surface $2. 10^{-5}$ cm^2. Etant donné que le diagramme de rayonnement émis par la face supérieure du cristal est lambertien, le coefficient de couplage dans la fibre, $n^2\Delta n$, peut être de l'ordre de 3%. Pour des courants d'injection respectifs de 6 mA et 60 mA le tableau 2 montre que la puissance lumineuse totale émise est de 100 µW et 1 mW, la puissance lumineuse introduite dans la fibre est 3 µW et 30 µW. Cette dernière valeur parait compatible avec des débits d'information de 10 Mbits/s transmis sur des distances de quelques km sans répéteurs.

Structure	Ø (μ)	Ø (cm^2)	I (mA)	J (A/cm^2)	P/I (mW/A)	B $W/cm^2/sr$	P (mW)	$P_0^{(1)}$ (μW)
Plane S.H	50	2.10^{-5}	6 60	300 3000	15 15	1,5 15	0,1 I	3 30
"de Burrus"	50	2.10^{-5}	200	10000	7-83	10-125	I-16	30-500
Laser	300×20		180	3000	75		12	6000

(1) avec $n^2\Delta n \sim 3\%$

Tableau 2:
Caractéristiques de la diode à confinement de porteurs (de structure plane, à simple hétérojonction) couplée à une fibre optique multimode: P_0 représente la puissance lumineuse insérée dans la fibre dans le cas d'un coefficient de couplage de 3%. Bien que la structure "de Burrus" ne satisfasse pas les conditions de stabilité, ses caractéristiques, ainsi que celle du laser sont rapportées par comparaison.

▪ IV. CONCLUSION

Nous avons montré que, dans l'état actuel des techniques, une durée d'utilisation garantie supérieure à plusieurs dizaines de milliers d'heures impose, outre certaines conditions relatives à la structure, une densité de courant modérée.

Considérant l'application des télécommunications optiques par fibre multimode, les diodes électroluminescentes apparaissent limitées d'une part par le produit courant d'excitation faible x efficacité quantique externe limitée aux alentours de 1%, d'autre part et tout autant par le diagramme spatial en loi de Lambert qui se traduit dans les conditions pratiques par un coefficient de couplage aux fibres de l'ordre de quelques pour cents.

Cependant un type de source, la diode à confinement de porteurs apparait, par sa structure et ses propriétés, particulièrement adaptée à la transmission d'information à un débit de l'ordre de 10 Mbits/s. L'effort de recherche et de développement doit se poursuivre en premier lieu sur la nature et la prévention des mécanismes de dégradation des sources à semiconducteurs. Ceci est une condition nécessaire à l'introduction des lasers et à l'élargissement des domaines d'application des diodes électroluminescentes.

Remerciements

L'étude des mécanismes de dégradation des diodes électroluminescentes et des moyens de prévention de ceux-ci a bénéficié à plusieurs reprises d'un soutien de la DGRST. La diode à confinement de porteurs de structure plane est en cours d'étude avec le soutien du CNET; l'épitaxie liquide de GaAlAs est étudiée avec l'aide de la DRME.

Ces études sont effectuées avec la collaboration de E. André, D. Diguet et J.C. Dubois.

REFERENCES

C.A. Burrus, Proc. of the IEEE, 60,231,1972.

J. Lebailly, C. Gouin, A. Desombre,10th annual proceedings of the IEEE Reliability Physics Symposium, 126,1972.

J. R. Biard, G.E. Pittman, J.F. Leezer, 1956 Symposium on GaAs.

J. C. Dubois, C. Gouin, J. Lebailly, résultats internes non publiés.

W.A. Brantley, D.A. Harrison, 11th annual proceedings on the IEEE Reliability Physics Symposium, 267,1973.

R.L. Hartman, M. Kuhn, 10th annual proceedings on the IEEE Reliability Physics Symposium, 137,1972.

G.A. Rozgonpi, C.J. Hwang, T.T. Ciesielka, J. Electochem. Soc. 120,333,1973.

J. Lebailly, D. Diguet, Journal de Physique, 35,C3-233,1974.

PHOTON STATISTICS

F.T. Arecchi

Università di Pavia and CISE

Milano, Italy

I. RELEVANCE OF PHOTON STATISTICS (PS)

Consider a photodetector illuminated by a light beam. By an electronic gate lasting for a time T, one counts the number n of photons annihilated at the photosurface in T. The random variable n has a statistical distribution p(n) that can be determined by iterating the above procedure for a large number of samples.

In Figure 1 we give an experimental plot of the statistical distribution of photocounts p(n) versus the number of counts

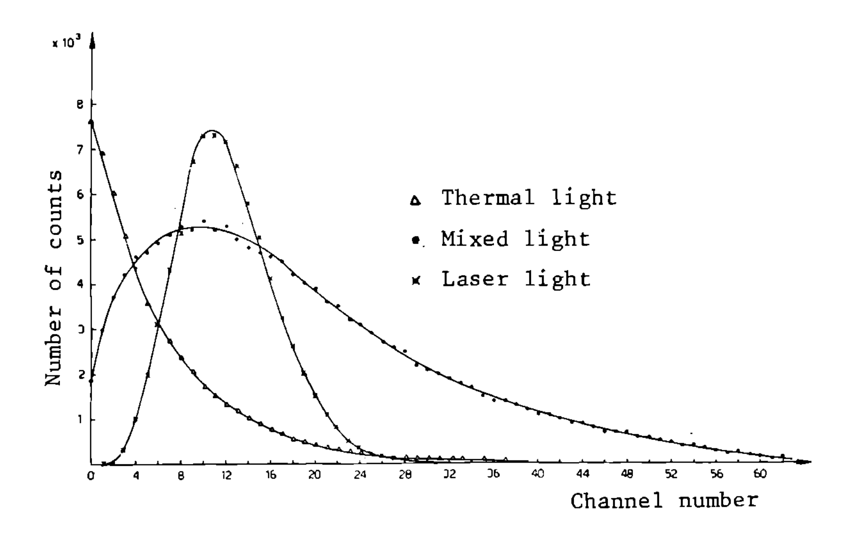


Figure 1 :
Photocount distributions

$$n = \eta \langle I \rangle T$$

(<I> being the average intensity, T the gating time ,η a constant accounting for the quantum efficiency of the detector plus other instrumental factors). The three curves refer to three physical cases which are indistinguishable from the point of view of classical optics. Indeed in the three cases we have the same average photon number $\langle n \rangle$, the same diffraction limited plane wave, the same linewidth $\Delta\omega$ filtered out in such a way that

$$\tau = 1/\Delta\omega >> T$$

that is, each sample of the statistical distribution p(n) is collected over a time T during which the field amplitude $|E|$ is practically constant. From the point of view of PS, the three lights are dramatically different as seen from the figure. The variances associated with distributions a) and b) are respectively :

$$\langle \Delta n^2 \rangle = \langle n \rangle \qquad (1.2)$$

$$\langle \Delta n^2 \rangle = \langle n \rangle + \langle n \rangle^2$$

In the laser case p(n) is fitted by the familiar Poisson distribution which describes the number fluctuation in a volume containing a classical gas in equilibrium. The relative r.m.s. fluctuation is

$$\frac{\langle \Delta n^2 \rangle^{1/2}}{\langle n \rangle} = \frac{1}{\sqrt{\langle n \rangle}}$$

and for $\langle n \rangle >> 1$ becomes negligible, justifying a description in terms of averages , as done in thermodynamics.

In the thermal case (light source in thermal equilibrium as for the black-body) the relative r.m.s. fluctuation does not decrease

$$\frac{\langle \Delta n^2 \rangle^{1/2}}{\langle n \rangle} \sim 1$$

hence, it is misleading to describe the field only in terms of average values $\langle n \rangle$.

The first contribution of (1.2) is a particle-like noise, as in the Poisson case (1.1) ; the second is a wave-like noise. If now one superposes a laser field with average photon number S and a thermal field with average $\langle n \rangle$ onto the same space mode, one obtains the sum of the two variances plus an interference term

$$\langle \Delta n^2 \rangle = S + \langle n \rangle + \langle n^2 \rangle + 2S \langle n \rangle \qquad (1.3)$$

To show the importance of the last term consider a communication channel with S coherent photons and $\langle n \rangle$ noisy thermal photons. If e.g. $S = 10^4$, we get for the different $\langle n \rangle$ values :

$\langle n \rangle$	$\langle \Delta n^2 \rangle$
0	10^4
10	20×10^4

This shows the practical importance of these statistical considerations.

■ II LIMITS OF CLASSICAL OPTICS

We show in this Section why the above questions are not accounted for in classical optics. We can expand the free field in a given region of space and time in orthonormal modes

$$\vec{E}(r,t) = \vec{E}^{(+)}(r,t) + \vec{E}^{(-)}(r,t)$$

$$= \sum_K [C_K \vec{u}_K(r) e^{-i\omega_K t} + C_K^\star \vec{u}_K^\star(r) e^{i\omega_K t}] \qquad (2.1)$$

where $E^{(+)}$, $E^{(-)}$ denote the positive and negative frequency parts of the Fourier expansion, and $u_K(r)$ can be calculated by suitable boundary conditions.

We express our ignorance on the sources of the field by saying that the complex field amplitudes C_K are random quantities assigned through a probability distribution

$$p(\{C_K\}) = p(C_1, C_2, \ldots C_K, \ldots) \qquad (2.2)$$

so that any field function is given in average by :

$$\langle f(E) \rangle = \int f[E^{(+)}(\{C_K\})]\, p(\{C_K\}) \prod_K d^2C_K + \text{c.c} \qquad (2.3)$$

where :

$$d^2C_K = d(\text{Re } C_K)\, d(\text{Im } C_K)$$

We see therefore that a *complete* characterization of the field (2.1) implies knowledge of the joint probability of *all* the amplitudes C_K or, equivalently, of the correlation functions of the fields

$$G^{(n,m)} = \langle E^{+}(X_1)E^{(+)}(X_2) \ldots E^{(+)}(X_n)\, E^{(-)}(X_{n+1}) \ldots E^{(-)}(X_{n+m}) \rangle$$

$$(X_i \equiv r_i \,,\, t_i) \qquad (2.4)$$

at any order (the ensemble average being performed as in (2.3)).

Directionality and monochromaticity, the properties associated with classical interferometers, give instead information on the *lowest order* correlation function , namely

$$G^{(1,1)} = \langle E^{(+)} (r_1 \; t_1) \; E^{(-)} (r_2 \; t_2) \rangle$$

This can be shown with reference to the Young and Michelson interferometers (Figure 2 and 3).

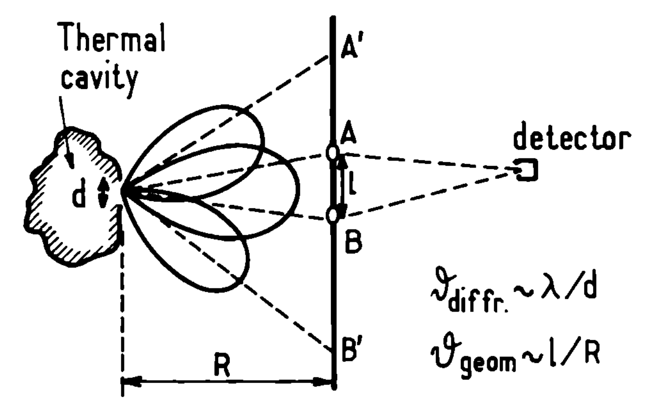


Figure 2 :
Young interferometer and coherence area. To see fringes, geometrical angle ℓ/R must be smaller than diffraction angle λ/d, hence (for solid angles) $\ell^2 < (R\ \lambda/d)^2 \equiv A$ (coherence)

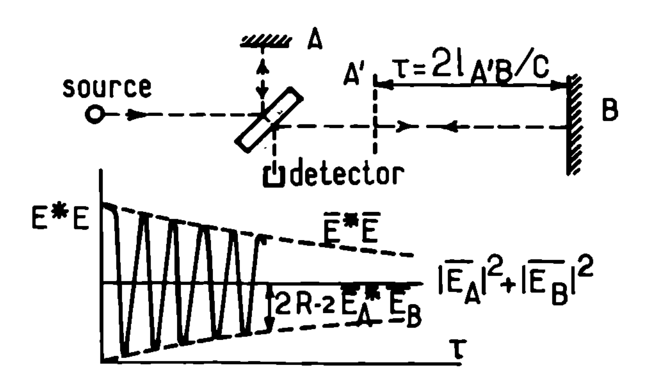


Figure 3 :
Michelson interferometer and coherence time. To see fringes, τ must be smaller than the reciprocal of the source linewidth.

Let us consider a radiating cavity with an aperture d(Figure 2) the expansion (2.1) inside that cavity each outgoing plane wave is broadened by diffraction by a solid angle $\theta^2_d \sim (\lambda/d)^2$. In Figure 2 we sketch for each plane wave a polar diagram, centered at the center of the aperture, giving the distribution of the intensities around the central propagation vector k for three different plane waves. The Young interferometer is made of a screen with two diffraction holes A,B plus a square-law detector which time-averages the square of the instantaneous signal

$$| E_A + E_B |^2$$

over a resolving time much longer than an optical period. As the detector moves parallel to the screen, it collects an average contribution $\overline{|E_A|^2} + \overline{|E_B|^2}$(bar means time average) plus a cross-term contribution $2\mathrm{Re}\,\overline{E_A E_B}$ which can be either positive or negative depending on the phase relationship between E_A and E_B. This extra contribution gives rise to fringes.

If the source is in thermal equilibrium, we expect that the plane waves of the expansion (2.1) are mutually independent. Therefore the cross-term is non-zero only when the holes A,B pick up fields coming from the same diffracted plane wave, that is, fringes appear only when the distance between A and B is such that the geometrical aperture d/R is smaller than the diffraction aperture λ/d (Figure 2).

The fringes are an indication of selection of a single k (directionality).

Hence directionality is related to the quantity.

$$G^{(1,1)}(r_A, r_B) = \langle E^{(+)}(r_A,t)\, E^{(-)}(r_B,t) \rangle \qquad (2.8)$$

Similarly, in a Michelson interferometer (Figure 3) one investigates the monochromaticity of a source of bandwidth $\Delta\omega$ by introducing a delay

$$\tau = t_B - t_A$$

between the two mirrors B and A, and looking at the fringe decay as τ is increased. Here light is correlated at the *same space* point but at different times, and the cross term will be proportional to

$$\overline{E^{(+)}(r,t_A)E^{(-)}(r,t_B)} = \sum_{\Delta\omega} {}_{\ell m}\, C_\ell\, e^{-i\omega_\ell t}\, \overline{C^*_m\, e^{i\omega_m}} (t-\tau)$$

we have limited the sum over the linewidth $\Delta\omega$. Replacing the time average with an ensemble average, and using the assumed mode independence it results $\langle C_\ell\, C^*_m \rangle = |C_\ell|^2 \delta_{\ell m}$. Hence the sum reduces to

$$\sum_\ell |C_\ell|^2\, e^{-i\omega_\ell \tau} \sim |C_0|^2 \sum_{\Delta\omega} e^{-i\omega_\ell \tau}$$

and it gives a non-zero contribution only when the complex numbers

are almost phased, that is up to $\tau|\omega_\ell - \omega_{\ell'}| \lessapprox 1$ or

$$\tau \approx 1/\Delta\omega$$

To conclude, the decay of fringes is associated with the decay of the term

$$G^{(1,1)}(t_A, t_B) = \langle E^{(+)}(r, t_A)\, E^{(-)}(r, t_B)\rangle \qquad (2.6)$$

In both cases, (2.5) and (2.6) give the number of k values or ω values within the resolution range of the interferometer, i.e. they show how to select a single mode from the mode expansion (2.1). We still have to solve the problem : how is each mode statistically distributed ? We must go over to higher order correlations as (2.4) with $n,m > 1$.

▪ III. CHARACTERIZATION OF RANDOM PROCESSES.

A random process is described by a random function $y(t)$ of time, which in realistic experiments we take as proceeding by quantized steps τ_R. (τ_R = resolving time of the measuring set-up) and varying over a finite interval $(0,T_0)$. If we consider n values $t_1, t_2 \dots t_n$ of t in $(0,T_0)$, then the values $y(t_1), y(t_2) \dots y(t_n)$ are a set of random variables characterized by the hierarchy of joint probability distributions

$$W_1(y_1) \qquad (3.1)$$

(probability density of finding $y(t_1)$ between y_1 and $y_1 + dy_1$ normalized as

$$\int_{-\infty}^{\infty} W_1(y_1)\, dy_1 = 1 \;)$$

$$W_2(y_1, y_2) \equiv W_1(y_1)\, P_c(y_1 | y_2) \qquad (3.2)$$

(joint probability of finding y_1 at t_1 and y_2 at t_2 ; this defines also a conditional probability of finding y_2 at t_2, given y_1 at t_1 through (3.2)) and so on. One should go up to W_n to fully characterize the process in $(0,T_0)$. However there are two simple random processes. The first is the fully random process in which no memory is kept of previous events, hence

$$W_n(y_1, y_2 \dots y_n) = W_1(y_1)\, W_1(y_2) \dots W_1(y_n)$$

The second is the Markoff process, completely characterized by W_2 , in the sense that the conditional probability is with "short memory" as expressed by the self-explanatory relation

$$W_3(y_1\, y_2\, y_3) = W_2(y_1\, y_2) P_c(y_1\, y_2 \,|\, y_3)$$
$$= W_2(y_1\, y_2)\, P_c(y_2 \,|\, y_3)$$

that is, P_c has no memory of y_1 but only of y_2 .

An equivalent way of describing a random process is through the correlation functions

$$< y_1 \, y_2 \, \cdots \, y_n > \equiv \int y_1 \, y_2 \cdots y_n \; W_n(y_1 y_2 y_n) \, dy_1 \, dy_2 \ldots dy_n \quad (3.3)$$

for any set of times $(t_1 t_2 \ldots t_n)$ including repeated indices. When the random process is an electric field, of particular importance are the first two correlation functions ,

$$G^{(1)} (1,2) \equiv < E_1^{\star} \; E_2 > \qquad (3.4)$$

which in particular for $t_1 = t_2$ becomes

$$I \equiv < E_1^{\star} \; E_1 > \qquad (3.4)^1$$

and $$G^{(2)}(1,2) \equiv < I_1 \; I_2 > = < E_1^{\star} E_1 E_2^{\star} E_2 > \qquad (3.5)$$

The field correlation function was used in Sec. 2. The intensity correlation function will be used in Sec 4.

▪ IV. GAUSSIAN PROCESSES AND THE HANBURY-BROWN AND TWISS EFFECT

Gaussian processes with zero average are those whose W_n are Gaussian :

$$W_1 (y) = N \, e^{-y^2/\sigma^2} \qquad (4.1)$$

and so on. An equivalent definition is by saying that the odd correlation functions are zero and the even ones factor out in all products of pairs :

$$\begin{aligned} &< y_1 \, y_2 \cdots y_{2n+1} > = 0 \\ &< y_1 \, y_2 \cdots y_{2n} > = \sum_P < y_1 y_2 > < y_3 \, y_4 > \cdots\cdots < y_{2n-1} \; y_{2n} > \end{aligned} \qquad (4.2)$$

where $\sum_P$ is sum over all permutations of (1.... 2n).

For example :

$$\begin{aligned} < y_1 y_2 y_3 y_4 > &= < y_1 y_2 > < y_3 \, y_4 > \\ &+ < y_1 y_3 > < y_2 \, y_4 > \\ &+ < y_1 y_4 > < y_2 \, y_3 > \end{aligned} \qquad (4.3)$$

For a complex field, combining (3.5) and (4.3) one gets

$$\langle I_1 I_2 \rangle = \langle I_1 \rangle \langle I_2 \rangle + |\langle E_1^* E_2 \rangle|^2 \qquad (4.4)$$

This is a very important relation, showing that the Gaussian intensity correlation sheds information on the square of the field correlation (Figure 4).

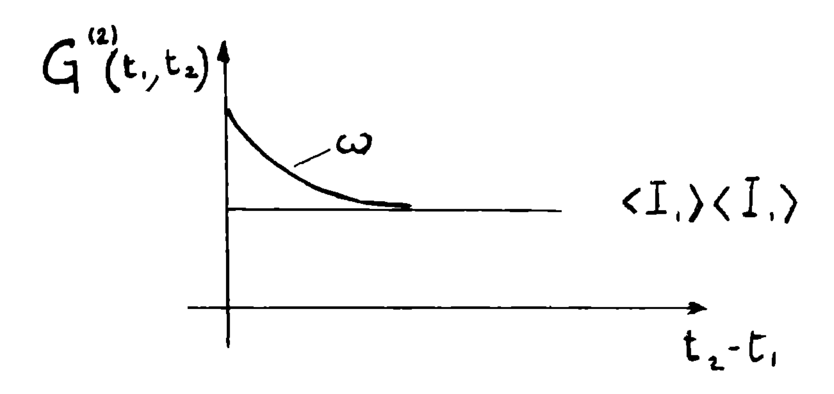


Figure 4 :
Plot of the intensity correlation function.

Gaussian processes are important for two reasons.

A) Entropy argument

In a cavity at thermal equilibrium the entropy as a function of the variable E is a maximum S_0 ; hence expanding around S_0 one has

$$S(E) = S_0 - \alpha E^2$$

But the field probability is given by Boltzman relation

$$W(E) = e^{S/k} \to e^{S_0/k} \quad e^{-(\alpha/k)E^2} \quad ,$$

hence it is Gaussian

B) Central limit argument

By the central limit theorem, the sum of many uncorrelated events is Gaussian. Such is the case of light generated by the atomic spontaneous emissions in a thermal source, or the light scattered by microscopic bodies in a light scattering experiment. In such a case, the scattered field E_s is proportional to the impinging field E_0 and to the random polarizability $\Delta\varepsilon$ of the scattering medium :

$$E_s \propto \Delta\varepsilon . E_0$$

If E_0 is free from fluctuations (say, laser light) then the correlations in E_s repeat faithfully the correlations in $\Delta\varepsilon$ and hence give information on the medium behaviour. Eq. (4.4) says that the *intensity correlation* (as obtained with a photodetector plus an electronic correlator) gives *all* information on the scatterer (except for frequency shifts). Historically , all that was born by an extrapolation of the interferometer idea.

In 1956 it was introduced by Hanbury-Brown and Twiss a new interferometer correlating the outputs of two detectors, and therefore correlating the intensities rather than the fields (Figure 5).

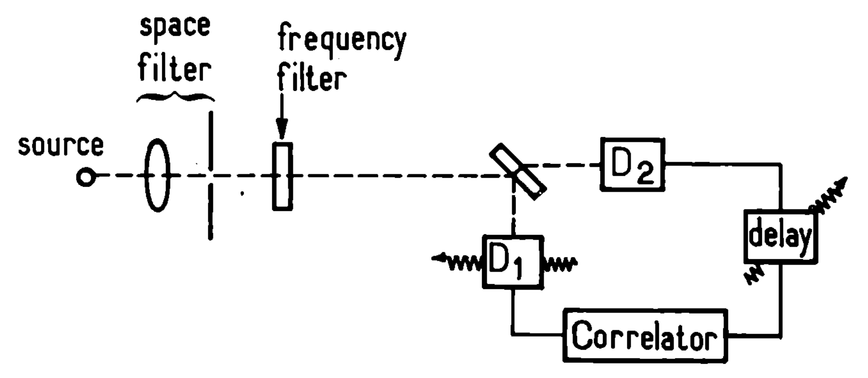


Figure 5 :
Hanbury-Brown-Twiss interferometer D_I, D_2 *detectors. Movement in* D_I *to pick up different angular regions ; variable delay between* D_I *and* D_2 .

The outcome of the experiment is here proportional to :

$$G(2,2) = \langle |E_1|^2 \; |E_2|^2 \rangle \qquad (4.5)$$

where E_1, E_2 are the fields at the two detectors. Suppose we have already selected by filters a single monochromatic plane wave :

$$E(r,t) = C\, e^{-i(\omega t - kr)} + C^{\star}\, e^{i(\omega t - kr)}$$

Then (4.5) this time implies the statistical distribution of C, *not only* its average intensity, since it is proportional to the fourth moment :

$$\langle |C|^2 \; |C|^2 \rangle = \langle |C|^4 \rangle$$

Two typical examples are :

1. *Gaussian distribution with zero average :*

$$p(C) = \frac{1}{\pi<n>} e^{-|C|^2/<n>} \qquad (4.6)$$

(4.6) corresponds to thermal equilibrium. Indeed it is a Boltzmann distribution $e^{-W/kT}/Z$ with $W= n\hbar\omega$ (n = photon number $\propto |C|^2$). It is well known that a complex Gaussian distribution as (4.6) has the following relation between fourth and second moment (see (4.4)) :

$$<|C|^4> = 2<|C|^2> \qquad (4.7)$$

2. *"coherent" field without fluctuations :*

$$p(C) = \delta^{(2)} (C - C_0) \qquad (4.8)$$

($\delta^{(2)}$ is a Dirac δ -function in the complex plane G)

In such a case the intensity correlation is

$$< |C|^4 > = < |C|^2 >= | C_0 |^4 \qquad (4.9)$$

Equations (4.8) and (4.9) suggest that coherence can be defined as : "δ-like amplitude distribution", or : "correlation functions factorized at any order".[3]

The above equivalent definitions can be generalized to include a many-mode field. A field made of, say, two modes is coherent when :

$$p(C_1 , C_2)= \delta^{(2)} (C_1 - C_{10}) \delta^{(2)} (C_2 - C_{20}) \qquad (4.10)$$

or

$$< E_1^{(+)} E_2^{(+)} .... E_{n+m}^{(-)} > = < E_1^{(+)}> < E_2^{(+)} > ..... < E_{n+m}^{(-)} > \qquad (4.11)$$

but it may be neither monochromatic nor unidirectional !

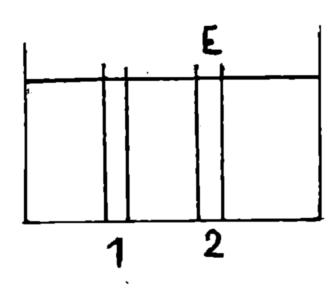
▪ V. MEASUREMENT OF THE PS

We give here a simple-minded picture based on the argument that photons being particles with zero mass cannot be localized when the field is uniform (Figure 6a).

Hence there are no a priori correlation between the outputs of two detectors 1 and 2, and the photons, whose average number is proportional to the square field and the measuring time T

$$< n > = |E|^2 T \eta$$

(η = quantum efficiency of the detector)



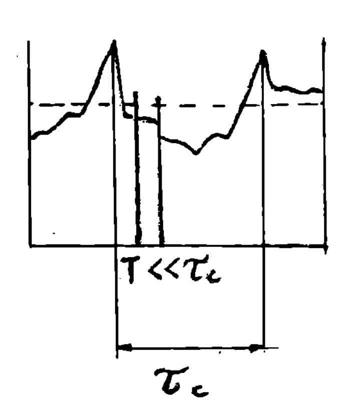


Figure 6 :
Profile of the field amplitude for a coherent and a fluctuation field case. τ_c *= coherence time (decay time of the field correlation function) ; T = observation time.*

must be distributed as a Poissonian (as the statisticsof radioactive counts), that is :

$$p(n) \equiv K(E,\ T\ |n) = \frac{\langle n \rangle}{n!}\ e^{-\langle n \rangle} \qquad (5.1)$$

If now (Figure 6 b) the field is randomly distributed with a statistics W_1 (E,t) and each measurement lasts for a time T much smaller than the coherence time τ_c ,

$$T \ll \tau_c \quad ,$$

In order to have a constant field within each sample, then we must average the detector statistics (5.1) over the field statistics

$$p(n,\ T,t) = \int K(E,T|\ n)\quad W_1(E)\ dE$$

In Figure 7 the results are shown pictorially for the three cases of Figure 1.

The photodetector used in measurements is a high-gain low-noise multiplier phototube. Anode-current pulses corresponding to single photoelectrons are standardized in amplitude and shape by a nonlinear circuit. This way , we get rid of the amplitude fluctuations in the multiplication process on the dynodes. An integrating capacitor acts as a number to voltage converter, and its voltage suitably amplified is then classified by a multichannel pulse-height analyzer. An alternative way of counting the number of pulses in a given time interval T is to use a fast electronic scale, gated "on" for a time T, which records the number of pulses and is directly connected with the memory of the multichannel analyzer : this avoids the double conversion

process, thus increasing the rate at which counts can be accumulated.

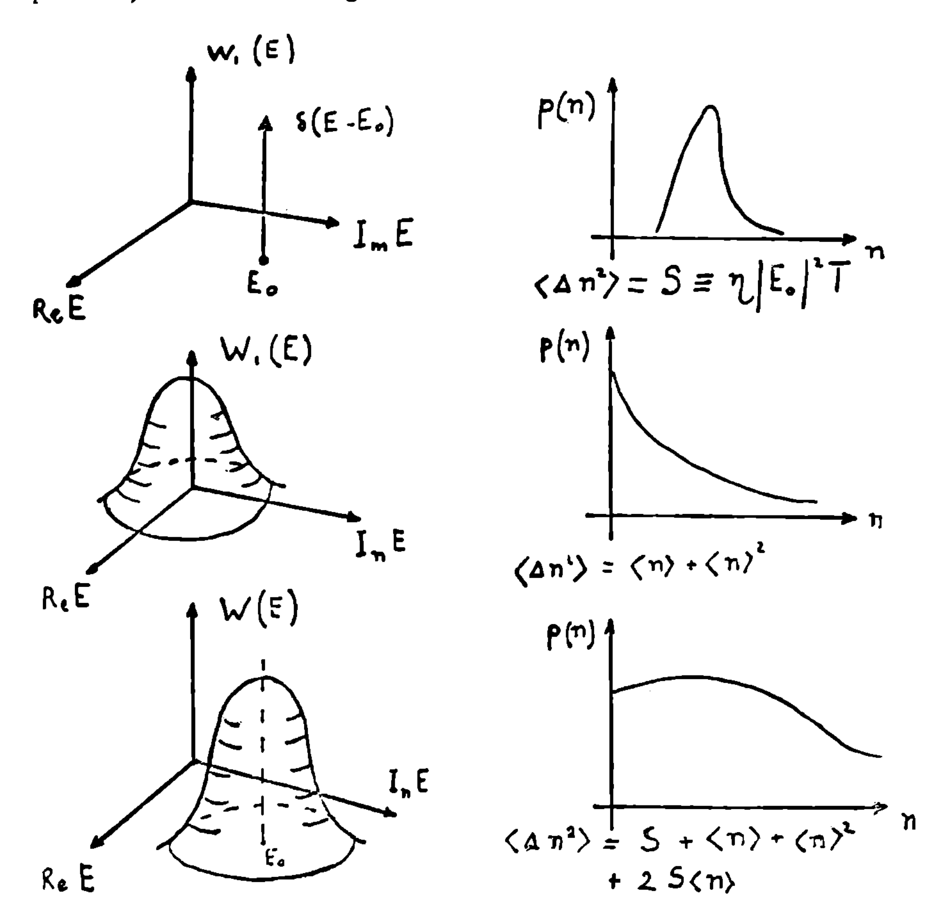


Figure 7 :
Qualitative plot of the statistical distributions of field and photons for the three cases of Figure I.

One obtains directly the distribution p(n,T) of photoelectron numbers. The distributions reported in Figure 1 have been obtained by this method.

A single p(n,T) gives only an integrated information on the time evolution of the field. One way of measuring the time evolution would be to measure the PS for increasing time-interval T up to (or larger than) the characteristic relaxation time of the field, and then to correlate the various shapes of the photocount distributions.

It is better however for the interpretation of the results to correlate separate observations, each one made for a time T much

shorter than the the coherence time, as shown in the experimental set-up (Figure 8).

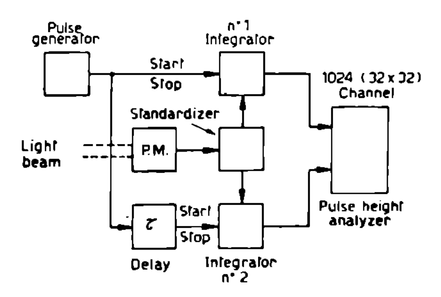


Figure 8 :
Experimental setup used for measuring joint photocount distributions

Essentially, the operation described before is repeated twice at times t_1 and t_2 and the two results are sent to a two-dimensional multichannel pulse-height analyzer. The results are classified on a two-dimensional matrix, which gives the joint photocount distribution

$$W_2(n_1t_1,\ n_2t_2) = P_c(n_1t_1 \mid n_2t_2)\ W_1(n_1t_1)$$

One easily realizes that P_c is given by the reading on a row defined by the chosen value of n_1) (Figure 9) and it is symmetrical with respect to indices 1 and 2 for all stationary processes. The marginal distribution $W_1(n_1) = W_1(n_2)$ corresponding to an uncorrelated experiment of the kind described previously, is obtained by summing for each column (row) the values corresponding to all rows (columns) belonging to that column(row) .

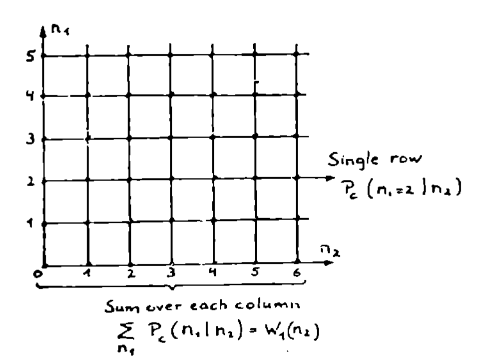


Figure 9 :
Matrix of the output numbers (each represented by a dot) as they are printed at the digital output of the mutichannel analyser. The numbers on a given line(e.g the row n_I=2) give a conditional distribution.

In Figure 10 we give the experimental results for a stationary Gaussian field together with the theoretical curves.

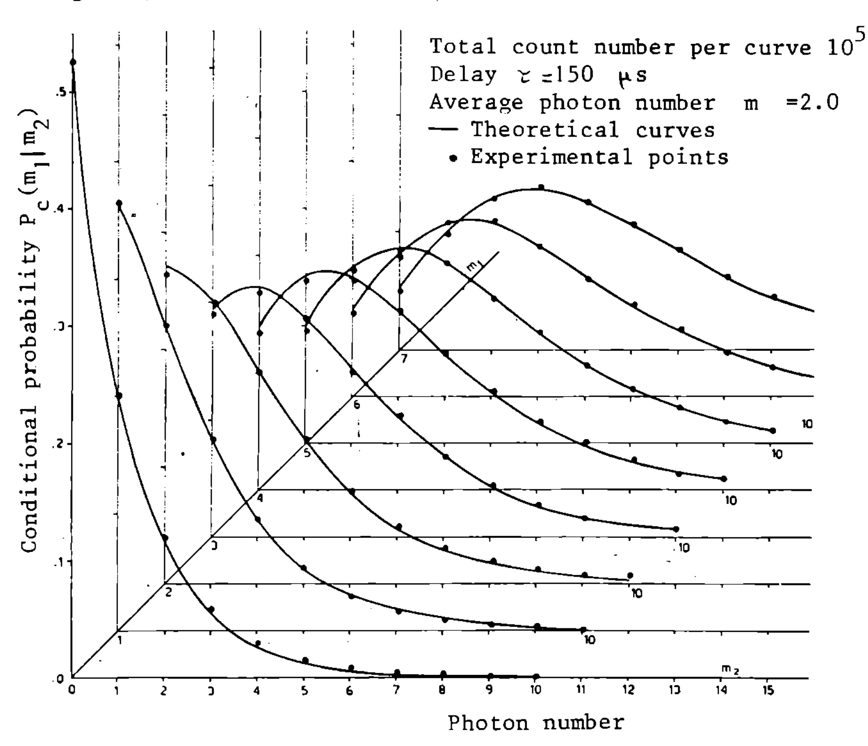


Figure 10
Joint photocount distribution for a gaussian field (laser light scattered by a rotating ground glass disk) with a coherence time of 700 μsec. The delay is I50 μsec (Arecchi et al I966 b).

VI. LASER FLUCTUATIONS

A) A review of the theory

As an application of the PS we present a set of experiments on both the stationary and transient statistical properties of a laser system. Let us make a heuristic description of a laser.

First let us consider a damped harmonic oscillator. Its statistical amplitude can be described by means of a Langevin equation, which we write in a rotating frame, (that is, after a transformation $\alpha \to \alpha e^{-i\omega t}$) as

$$\dot{\alpha} + \gamma\alpha = \Gamma(t) \tag{6.1}$$

Here $\alpha(t)$ is the complex amplitude of the oscillator, γ is the damping constant, and $\Gamma(t)$ is a random noise source. Let $\Gamma(t)$ be a

complex stationary Gaussian process, with zero average and δ-correlated in time :

$$< \Gamma(t) > = 0 \qquad (6.2)$$

$$< \Gamma^{*}(t) \; \Gamma(0) > = Q \; \delta(t)$$

In order to have the conditional probability $P(\alpha_0, 0 | \alpha, t)$ we must solve the two-dimensional Fokker-Plank equation associated with eq. (6.1) which is

$$\frac{\partial P}{\partial t} - \gamma \, \mathrm{div}_\alpha \, (\alpha P) = q \, \nabla_\alpha^2 \, P \qquad (6.3)$$

where $q = Q/4$. Its complete solution is (Figure 11).

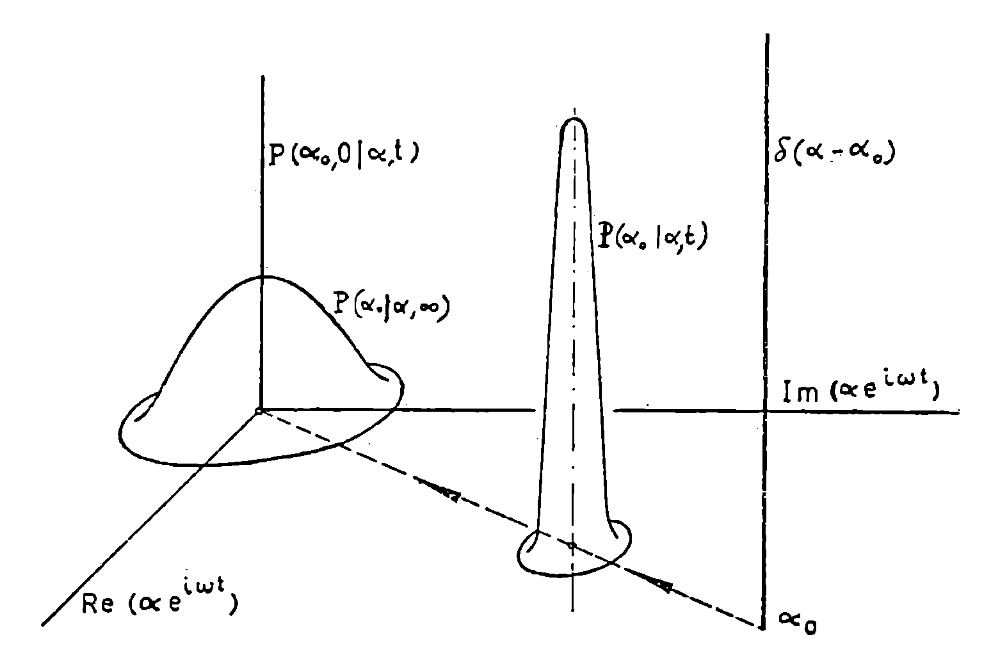


Figure 11 :
Evolution of the conditional probability for a damped harmonic oscillator.

$$P(\alpha_0, 0 \,|\, \alpha, t) = \frac{1}{\pi \sigma^2(t)} \, e^{- \frac{|\alpha - \alpha_0 \exp(-\gamma t)|^2}{\sigma^2(t)}} \qquad (6.4)$$

where

$$\sigma^2(t) = \frac{2q}{\gamma} \, (1 - e^{-2\gamma t})$$

In the limit $t \to \infty$ we obtain the first-order probability density

$$\lim_{t\to\infty} P(\alpha_0 0|\alpha t) = W_1(\alpha) = \frac{1}{\pi 2q/\gamma} e^{-\frac{|\alpha|^2}{2q/\gamma}} \tag{6.5}$$

By equation (6.3) the oscillator amplitude has been described as a stochastic Markoff process, whose ensemble distribution is a Gaussian with variance $< n>= 2q/\gamma$ and whose frequency spectrum is Lorentzian with half-width $\Delta\omega = \gamma$ (i.e whose first-order correlation function has an experimental time decay with time constant $1/\gamma$).

We may apply this treatment to a single-mode laser field and think of it as an oscillator whose damping is compensated by a gain term G due to the excited atoms in the cavity. The associated Langevin equation is

$$\dot{\alpha} + (\gamma - G)\alpha = \Gamma(t) \tag{6.6}$$

Generally G is amplitude dependent, but in first approximation it can be considered as a constant $G=g_0$. In this case the same solution (6.4), (6.5) still hold, in which we have simply to replace γ with $\gamma - \gamma g_0$. A more realistic picture is obtained considering the dependence of G on $|\alpha|^2$(cubic nonlinearity in the induced dipole moment), that is

$$G = g_0 - \beta |\alpha|^2 \tag{6.7}$$

The Langevin equation becomes

$$\dot{\alpha} - \beta(d - |\alpha|^2)\alpha = \Gamma(t) \tag{6.8}$$

where the linear negative damping $\beta d = g_0 - \gamma$ comes from the linear theory and $\beta|\alpha|^2$ is the first nonlinear correction. The Fokker-Plank equation becomes more complicated and we give only the final results. Its stationary solution in the plane $\alpha \equiv re^{i\phi}$ is independent of ϕ and is given by :

$$P(r) = Ne^{-\frac{\beta}{4q} r^4 + \frac{\beta d}{2q} r^2} \tag{6.9}$$

It is suitable to introduce the pumping parameter

$$a = \sqrt{\frac{\beta}{q}}\, d \tag{6.10}$$

and to normalize the square modulus as $X^2 = \sqrt{\frac{\beta}{q}}\, r^2$. The P(r) distribution is then written as

$$P(X) = Ne^{-\frac{1}{4}X^4 + \frac{a}{2}X^2} \tag{6.11}$$

We define the threshold point as $\gamma = g_0$, that is, as that point where the linear gain is equal to the losses. At threshold

$$d = a = 0$$

Below threshold ($a < 0$), the distribution (6.11) has a peak at $x=0$, whilst above threshold ($a > 0$) it exhibits a peak at $X = \sqrt{a}$. It is easily shown that, outside the interval around threshold, the following approximation hold :

1) Far below threshold ($a < -10$). Equation (6.11) reduces to :

$$P(r) \sim N e^{-\frac{|a|}{2} X^2} = N e^{-\frac{\beta|d|}{2q} r^2} \tag{6.12}$$

which is a Gaussian distribution centered at $r=0$, with a variance $< X^2 >$ equal to $2/|a|$ (Figure 14) for $a<0$).

2) Far above threshold ($a > 10$) .

The saturation becomes strong and tends to stabilize the field amplitude around its average value $< x > = \sqrt{a}$

The field statistics is

$$p(X) = N' e^{-a(X-\sqrt{a})^2} \tag{6.13}$$

This is a Gaussian distribution centered at $x = \sqrt{a}$.

$P(\alpha)$ as given in equation (6.13) is only a modulus distribution uniform in phase.

We give now the time evolution of the field in the same two limiting cases.

1) Far below threshold. The following formulas hold for the field and intensity correlation functions respectively :

$$G^{(1)}(\tau) = < \alpha^*(\tau) \; \alpha(0) > = \frac{2q}{\beta|d|} e^{-\beta|d|\tau} \tag{6.14}$$

$$G^{(2)}(\tau) = <|\alpha(\tau)|^2|\alpha(0)^2|> + \left(\frac{2q}{\beta|d|}\right)^2 e^{-2\beta|d|\tau} \tag{6.15}$$

The Fourier transforms of these expressions give the spectra of the field and intensity fluctuations. These spectra have Lorentzian shapes, half-widths $\lambda^{(1)} = \beta|d|$, $\lambda^{(2)} = 2\lambda^{(1)}$ respectively.

These behaviour are illustrated in Figures 12 and 13 for $a < 0$.

2) Far above threshold.

Far above threshold the field distribution is peaked at $\sqrt{d}$.

The laser field appears as the linear superposition of a coherent field with amplitude $\sqrt{d}$ plus a Gaussian field with zero average, photon number given by :

$$< n > = \frac{q}{2\beta d} \tag{6.16}$$

and decay constant increasing linearly with d, that is

$$\Delta\omega = 2\beta d \tag{6.17}$$

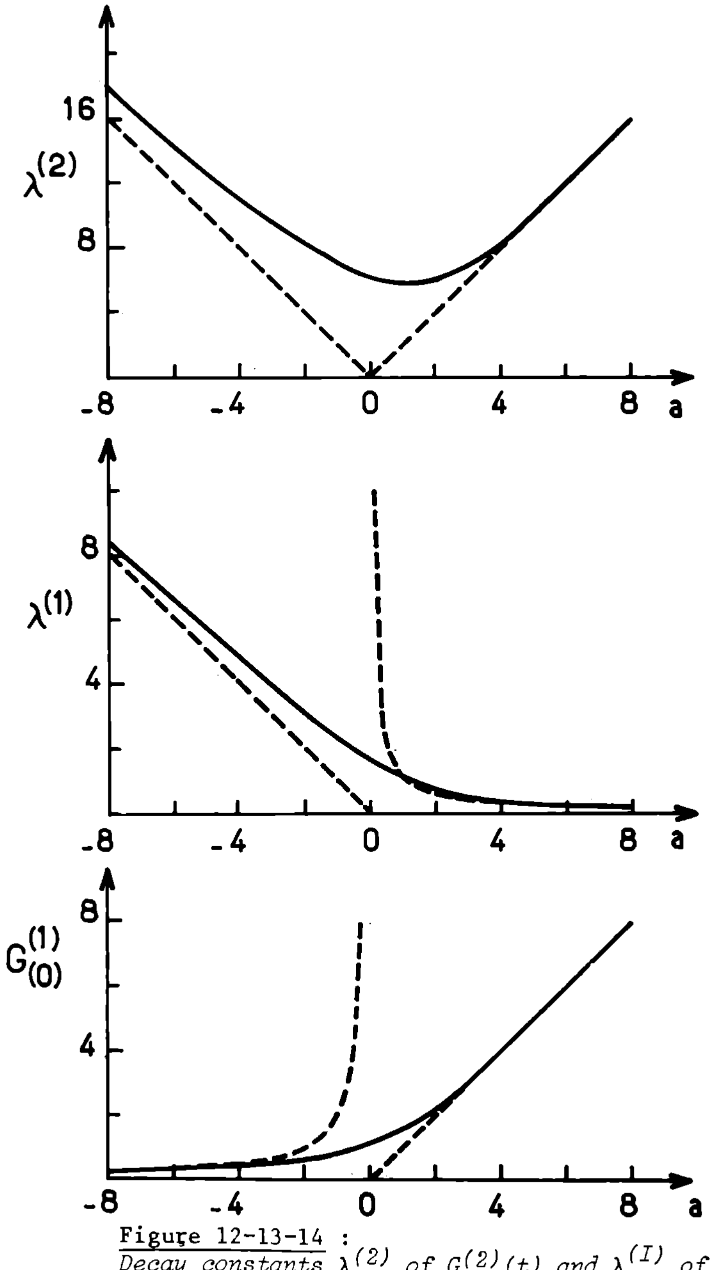


Figure 12-13-14 :
Decay constants $\lambda^{(2)}$ of $G^{(2)}(t)$ and $\lambda^{(I)}$ of $G^{(I)}$, and intensity $G^{(I)}(0)$, versus pump parameter a. Dashed lines : linearized theory. Solid lines : exact solution of Fokker-Planck equation.

As for the phase ϕ, it obeys a diffusion equation, whose solution has a diffusion constant q/d proportional to the reciprocal of the output power (Townes formula for a maser oscillator) ; see Figure 13 for $a > 0$.

3) Laser at threshold

So far we have discussed only the cases well above and below threshold. The behaviour in the threshold region cannot be obtained by simple extrapolation of the previous results, and we need a complete solution.

We refer to the theoretical calculations by Risken .[14] His solutions are plotted as solid lines in Figure 12 to 14 and are again reported for comparison in the experimental results (Figures 15 to 18). We see that the threshold point does not show discontinuities as it should appear from a naive linear picture.

B) Stationary experiments (ensemble distributions and time correlation). [12, 13]

In this Subsection we describe the experimental results obtained by means of the PS method in the study of the statistics of the e.m.field of a stabilized laser operating in different conditions. We used a 6328 Å He-Ne laser, single TEM_{00} mode, with one mirror supported by a piezoceramic disc in order to stabilize against fluctuations and to move the mode position with respect to the atomic line.

The measurement of p(n) is performed as described above. For comparing experiments and theory we used the second reduced factorial moment of the photocount distribution

$$H_2 = \frac{\langle n\,(n-1)\rangle}{\langle n\rangle^2} - 1 \qquad (6.18)$$

which goes from 1 (Gaussian field distribution ; well below threshold) to 0 for an amplitude-stabilized field (well above threshold), and the third one

$$H_3 = \frac{\langle n\,(n-1)\,(n-2)\rangle}{\langle n\rangle^3} - 1 \qquad (6.19)$$

which goes from 5 (Gaussian distribution) to 0 (amplitude-stabilized field) (Figure 15 and 16).

From the stationary solution of the Fokker-Planck equation for the statistical distribution of the laser field one can derive the distribution of photocounts and the associated factorial moments. One can see from the figures that the agreement between experiments and theory is very good. Finally we report the frequency spectra of the intensity fluctuations, and show that they are consistent with the time dependent solutions of the same Fokker-Planck equation, whose stationary solution is fitted by the ensemble distribution reported above.

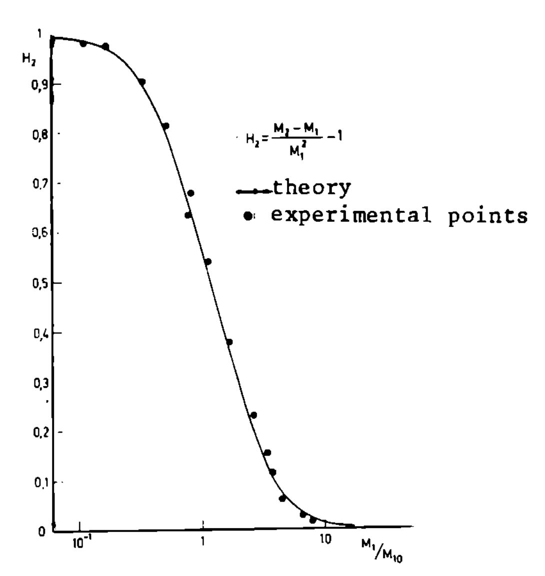


Figure 15 :
Measured and theoretical values of the reduced second-order factorial moment H_2 as a function of the normalized intensity M_I/M_{IO} in the threshold region.—Theory, • Experimental point.

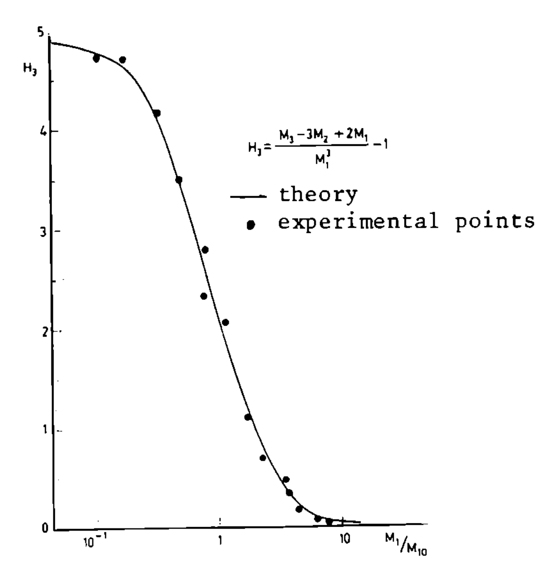


Figure 16 :
Measured and theoretical values of the reduced third-order factorial moment H_3 as a function of the normalized intensity M_I/M_{IO} in the threshold region.— Theory, • experimental point.

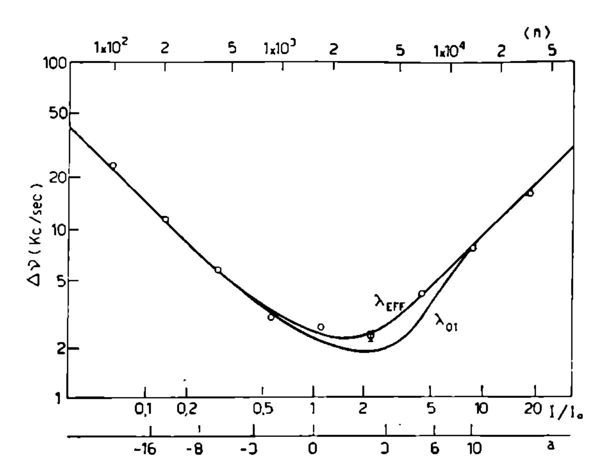


Figure 17 :
Plot of the "effective "linewidth $\Delta\nu$ *vs the laser intensity I normalized to the threshold values* I_0 . *The horizontal axis is also calibrated in values of the pump parameter a and in average photon number* $\langle n \rangle$ *inside the cavity.*

The results are reported in Figure 18 and 19.

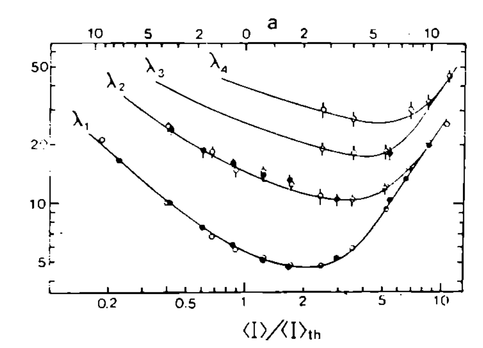


Figure 18 :
Plots of the relaxation rates of the laser intensity fluctuations. The full curves represent the theoretically predicted results. The standard deviations for the experimental points are reported only when they exceed the dot size. The two values of the proportionality constant C are 0,97 $\times 10^3$ *(open dots) and 1.03* $\times 10^3$ *(full dots).*

The power spectra of the intensity fluctuations measured with a wave analyzer showed small deviations from a Lorentzian shape in the threshold region. For any spectral shape, we define the "effective" linewidth $\Delta\nu$ of an equivalent Lorentzian as the ratio between

the total spectral power and the zero frequency value of the spectral density.

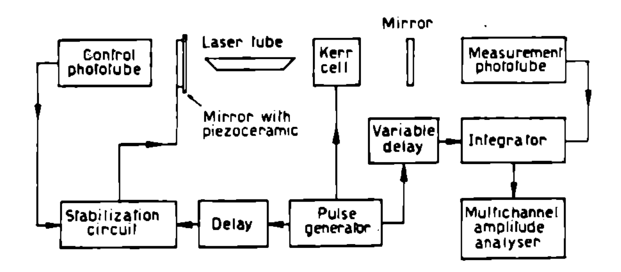


Figure 19 :
Experimental set-up for the transient experiment.

The measured linewidth fits the numerical calculations of Risken and Vollmer [14] who gave a dynamical solution for the intensity correlation function as a sum of several exponential terms (with decay constants $\lambda_k^{(2)}$) occuring with weights M_k. This leads to an equivalent decay constant.

$$\lambda_{eff}^{(2)} = \left(\sum_k \frac{M_k}{\lambda_k^{(2)}} \right)^{-1} \qquad \sum_k M_k = 1$$

In the Figure 18 we have also plotted the main decay constant of the intensity correlation function. It is clear that a Lorentzian approximation with a single decay constant is not an adequate description.

By using a correlator rather than a frequency analyzer it has been recently possible to measure the first four decay constant versus the laser intensity (Figure 18). [13 bis]

C) Transient experiments [15]

By joint use of a Q-switched gas laser and of the linear method for PS one can study a nonstationary statistical ensemble, measuring the time evolution of a laser field during a fast built-up.

The experimental set-up is shown in Figure 19. We put a Kerr cell with end faces at the Brewster angle within a single mode 6328Å He-Ne laser cavity.

Starting with some pre-set pump and cavity-parameters, but with the optical shutter closed, the Kerr cell is switched "on" in

a time shorter than 5 ns at the instant t=0. The laser field undergoes a transient build-up, from an initial statistical distribution corresponding to the equilibrium between gain and losses far below threshold up to an asymptotic condition above threshold. At the instant t= τ we perform photocount measurements for a measuring interval T of 50 ns, very small compared to the build-up time which is in our case of the order of some microseconds. Once a steady-state condition has been reached, an amplitude-stabilizing operation is performed by sampling the laser output and comparing this with a standard reference signal. This is equivalent to "preparing" an identical initial state for a successive measuring cycle.

After the sampling, the shutter is switched off for about 10ms. At the end of this interval the shutter is again switched on and the above described cycle of operations is repeated. This way, we collect an ensemble of macroscopically identical events. By varying successively τ , we obtain the time evolution of the photocount distribution p(n,T,).

A set of experimental results is shown in Figure 20.

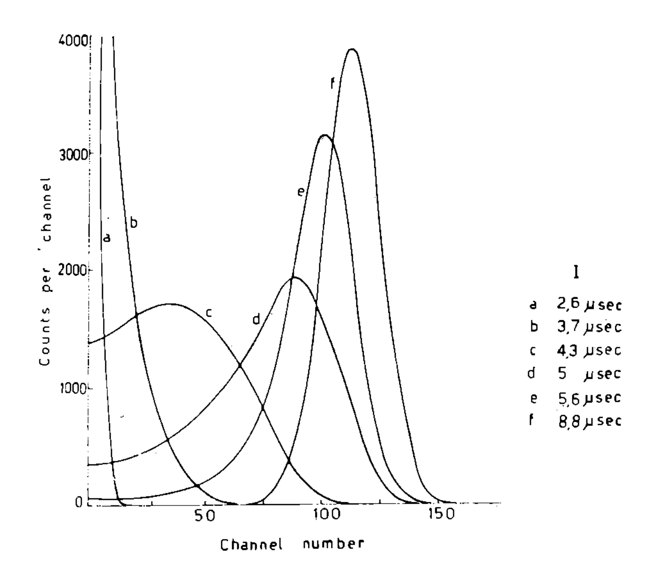


Figure 20 :
Experimental statistical distribution with different time delays obtained on a laser transient. The solid lines connect the experimental points which are not shown to make clearer the figure. All distributions are normalized to the same area a) 2.6μs; b) 3.7 μs; c) 4.3 μs; d) 5 μs; e) 5.6 μs; f) 8.8 μs.

The average photocount number $< n >$ and the associated variance $<\Delta n^2> = <n^2> - <n>^2$ are reported as function of the time delay in Figure 21 and 22 for two different pumping conditions.

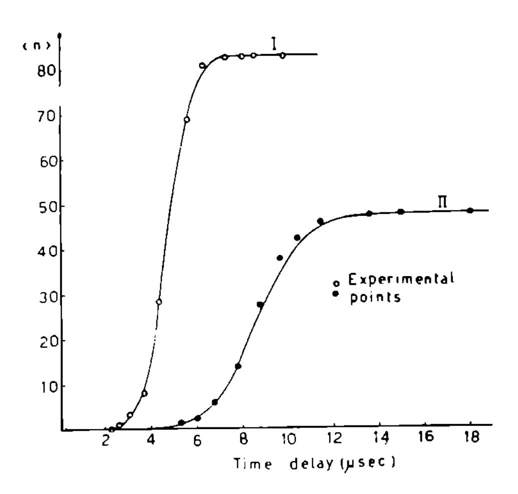


Figure 21 :
Evolution of the mean value $<n>$ of the statistical distribution $p(n,T,\tau)$. The solid lines represent the theoretical curves which best fit the experimental points(o, •).The experimental moints related to curve I) have been obtained by the statistical distribution of Fig 20.

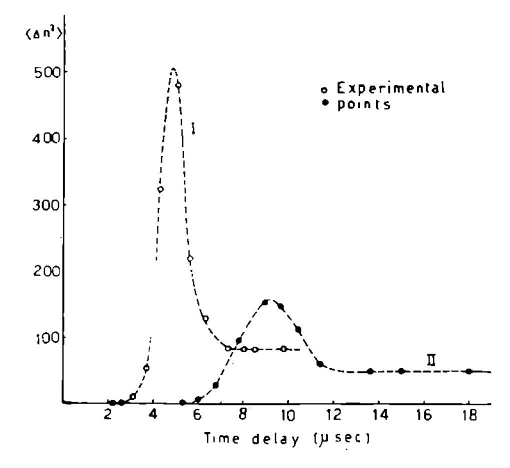


Figure 22 :
Evolution of the variance $<\Delta n^2> = <n^2> - <n>^2$ of the statistical distribution $p(n,T,\tau)$ Dashed lines represent an interpolation of the experimental points (c, •).

Notice that in Figure 22 the dashed line just interpolates the experimental points but has no theoretical significance. In Figure 23 we correct for attenuation, as described in the next Section, and we see that the experimental points agree with the theoretical, solid line.

VII. DISTORSION OF PS DUE TO ATTENUATION. [18]

The role of attenuation (Figure 24) is as follows.

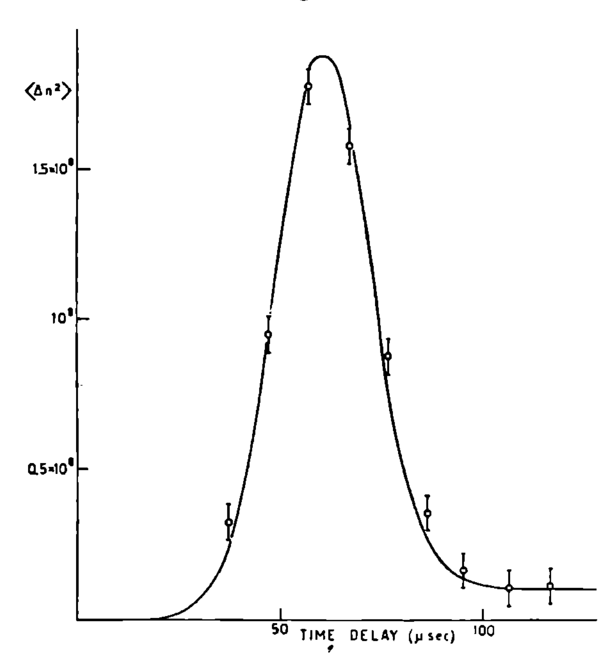


Figure 23 :
Evolution of the variance $\langle \Delta n^2 \rangle$ of the statistical distribution of photons inside the laser cavity, as a function of the time delay τ. The solid line represents theoretical results (Arrecchi and Degiorgio 1971).

For a n-photon field (as described in quantum mechanical books) the attenuation would give rise to a binomial statistics as the partition noise through the grid of an electron vacuum tube. However for a statistical mixture of coherent fields (each having a Poisson spread in photon number as described above) the attenuation is a purely classical process

$$E' = \eta E \tag{7.1}$$

which affects the statistics leaving unaffected the elementary probability

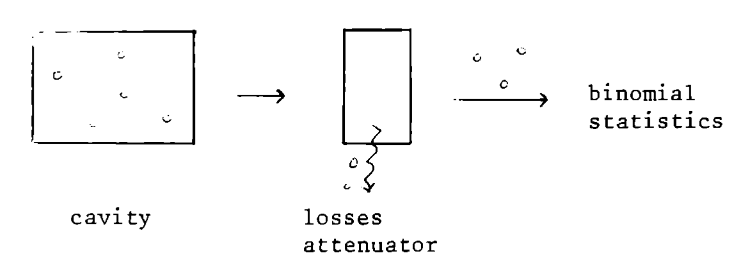
$$P(E')\, d^2E' = P(E)\, d^2E \tag{7.2}$$

By use of (7.2) it is easily shown that the factorial moments, given by :

$$F_k < n(n-1)\dots(n-k+1) > = \int |E|^{2k} P(E)\, d^2E \qquad (7.3)$$

change as

$$F'_k = \eta^{2k} F_k \qquad (7.4)$$



a) field with exactly "n" photons

E

E' = E

b) coherent field (or statistical mixture of coherent fields)

Figure 24 :

Picture of "particle-like" and "wave-like" attenuation.

By (7.4) one can perform corrections for an attenuation η. This kind of correction has led from Figures 22 to 23.

▪ VIII. THE PHOTOMULTIPLIER AS A STATISTICAL DEVICE.

The finite response time of the photodetector puts a limit on the use of PS for short times. We show here how intensity correlations can be corrected for the detector resolution. The output i(t) of a photocathode illuminated by a light signal may be represented by a train of photo-electrons localized at random times t_k, that is,

$$i(t) = \sum_{k=1} \delta(t - t_k) \qquad (8.1)$$

The first correlation function $R_i(\tau) = < i(t)\ i(t+\tau) >$ of the random process i(t) which we assume stationary, is given by :

$$\begin{aligned} R_i(\tau) &= <\sum_k \sum_j \delta(t - t_k)\ \delta(t + \tau - t_j)> \\ &= <\sum_k \delta(t-t_k)\ \delta(t+\tau-t_k)> + <\sum_{k\neq j} \sum_j \delta(t-t_k)\ \delta(t+\tau - t_j)> \qquad (8.2) \\ &= \mu\delta(\tau) + \mu^2 g^{(2)}(\tau), \end{aligned}$$

where μ is the average rate of photo-electrons, which is proportional to the average light intensity, and $g^{(2)}(\tau)$ is the reduced intensity correlation function.

The photomultiplier, considered as a linear stochastic filter (see Ref. [19]) is described by a random response function h(t) which is the output for a single photo-electron emitted at time t=0. Consequently the random output current u(t) will be the time convolution of the two random functions h(t) and i(t), that is,

$$u(t) = \int_{\infty}^{\infty} i(t-t')\ h(t')\ dt' \equiv i(t) \star h(t) \qquad (8.3)$$

In performing this convolution one should remember that the different $\delta(t-t_k)$ of the input (8.1) are associated with independent realizations $h_{(k)}(t)$ of the random process h(t). The correlation function $R_u(\tau) = < u(t)\ u(t+\tau) >$ is computed by taking into account that the statistical properties of the photomultiplier operating in the linear range are independent from the statistics of i(t) and that $h_{(k)}$ and $h_{(j)}$ for $k \neq j$ are statistically independent samples. It reads

$$R_u(\tau) = \mu^2\ g^{(2)}(\tau) \star G(\tau) + \mu C(\tau) \qquad (8.4)$$

where :

$$G(\tau) = \int_{\infty}^{\infty} <h(t)> < h(t+\tau)>\ dt \qquad (8.5)$$

and

$$C(\tau) = \int_{\infty}^{\infty} <h(t)\ \ h(t+\tau)> dt \qquad (8.6)$$

A full description of the effect of the photomultiplier on the intensity correlation function $R_u(\tau)$ requires therefore two different functions $C(\tau)$ and $G(\tau)$. $C(\tau)$ modifies the δ-like term of Equation (8.2) , which comes from the discrete character of the detection process, whereas $G(\tau)$ modifies the term describing the joint probability of two different photon events.

The fluctuations of h(t) are mostly generated by three effects : gain statistics at the dynodes, spread in transit times of secondary electrons from dynode to dynode, spread in the transit time of photo-electrons from the photocathode to the first dynode. If only the

first effects were present , $C(\tau)$ would be proportional to $G(\tau)$, whereas spreads in transit times affect the two functions differently. In particular, the spread associated with the transit time of photo-electrons affects only the location of each $h(t)$ and hence does not affect $C(\tau)$. As an example, we report in Figure 25 experimental results on both $C(\tau)$ and $G(\tau)$ (see Ref [19] and Ch. A5 of laser Hand-book). The two functions in Figure 25 are normalized to the same initial value.

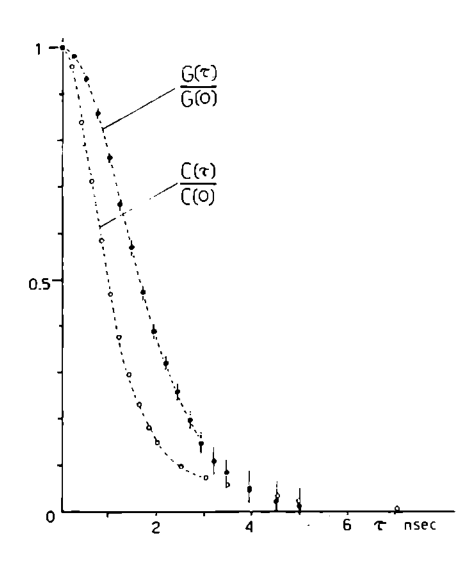


Figure 25 :

Plot of the functions C(τ) and G(τ) for a philips XP I2I0 photomultiplier, operating at 3800 V. The dashed lines interpolate the experimental points (Arecchi et Al I97I).

To obtain information on the light intensity correlation $g^{(2)}(\tau)$ from the measured correlation $R_u(\tau)$, the term $\mu C(\tau)$ must be subtracted. In this correction experimental errors strongly affect the results, especially at low photon flux. However the term $\mu C(\tau)$ does not appear at all in the expression for $R_u(\tau)$ when the correlation is performed with two phototubes. In that case, a similar calculation leads to a modified form of Equation (8.4).

$$R_u(\tau) = \langle u_1(t)\, u_2(t+\tau) \rangle = \mu_1\mu_2\, g^2(\tau) \star G_{12}(\tau) \qquad (8.7)$$

where the indices 1 and 2 refer to the two phototubes and

$$G_{12}(\tau) = \int_{-\infty}^{\infty} \langle h_1(t) \rangle \langle h_2(t+\tau) \rangle \, dt \qquad (8.8)$$

The treatment can be easily generalized to the measurement of

higher-order correlation functions.

■ REFERENCES

[1] R.J. Glauber : in *Quantum Optics and Electronics* (Proceedings of Les Houches 1964 Summer School) ed. by C.De Witt, et al. (New York 1965) plus previous references reported therein.

[2] R.J. Glauber : in Physics of Quantum Electronics (Proceedings of 1965 Puerto Rico Conference), ed. by P. Kelley, et al. (New York 1966).

[3] L. Mandel and E. Wolf : Rev. Mod. Phys. ,37, 213 (1965).

[4] M.C. Wang and G.E. Uhlenbeck : Rev. Mod. Phys.,17, 323 (1945).

[5] F.T. Arecchi : in Quantum Optics ed. by R.J. Glauber (1967 E. Fermi School)Academic Press.

[6] F.T. Arecchi : Phys. Rev. Lett. 15 , 912 (1965)

[7] F.T. Arecchi , A. Berné and P. Burlamacchi : Phys. Rev. Lett.16 ,32 (1966)

[8] F.T. Arecchi, A. Berné and A. Sona : Phys. Rev. Lett., 17 ,260 (1966)

[9] F.T. Arecchi , A. Berné, A. Sona and P. Burlamacchi: Proc of the IV International Quantum Electronics Conference, Phoenix 1966; IEEE Journ. Quantum Electronics, 2 , 341 (1966).

[10] F.T. Arecchi, E. Gatti and A. Sona : Phys. Lett. 20, 27 (1966)

[11] H. Risken : Zeits Phys.,186, 85 1965 ; 191 302 1966.

[12] F.T. Arecchi, G.S. Rodari and A. Sona :Phys.Lett. 25A ,59 1967.

[13] F.T. Arecchi, M. Giglio and A. Sona :Phys.Lett.25A ,341 1967.

[13] (Bis) M. Corti ,V. Degiorgio and F.T. Arecchi :Optics Commun, 8 329 (1973)

[14] H. Risken and H.D. Vollmer: Zeits. Phys. 201 ,323 (1967)

[15] F.T. Arecchi, V. Degiorgio and B. Querzola: Phys. Rev.Lett,19 ,1168 (1967); Arecchi-Degiorgio, Phys. Rev. A3 ,1108 (1971).

[16] W.E. Lamb: Phys.Rev. 134, A 1429 (1964)

[17] F.T. Arecchi ; M. Giglio and U. Tartari: Phys. Rev.163 (1967).

[18] F.T. Arecchi ,V. Degiorgio:Phys.Letters 27A ,429 (1968)

[19] F.T. Arecchi, M. Corti, V.Degiorgio, S.Donati :Opt. Comm.3, 284(1971)

Note : For a more up to date reference on Photon statistics and Laser fluctuations see Part A of

Laser Handbook (F.T. Arecchi, E.O. Schultz Du Bois eds). North Holland 1972.

PHOTODETECTORS

J. Belmas

Laboratoire de Marcoussis
Département Recherches Physique de Base
Section Composants Electroniques CGE
91460 Marcoussis France

■ I. INTRODUCTION

The field of applications of optical signals is growing rapidly, and the increase in the number of optical sources from the visible to the far infrared requires photodetectors which must be optimized from the point of view of their rapidity and sensitivity.

The purpose of this talk is to review the subject of semiconductor photodetector, which can be used in the spectral range from the visible to the near infrared.

■ II. BASIC PRINCIPLES

The absorption of optical radiation in a semiconductor material takes place by excitation of electrons and phonons and by the interactions between phonons and electrons.

We can distinguish five types of absorption in semiconductor materials ; for each of these types of absorption, the following effects occur :

- Electronic transitions between different energy bands
- Vibrations of the crystal lattice
- Electronic transitions within an energy band.
- Electronic transitions with the discrete levels of impurity atoms
- Vibrations of impurity atoms

In the range of wavelengths of interest, we only consider the band-to-band transitions, since the other effects do not occur or are of second order.

The band-to-band transitions will not occur unless the energy of the incident radiation is greater than, or on the order of the width of the band gap (Eg).

Radiation of a frequency ν less than Eg/h only exhibits losses by reflection, which depends on the refractive index of the semiconductor. This radiation is not absorbed until the frequency

becomes very low, equivalent to the characteristic lattice vibration.

However when the photon energy approaches Eg, the absorption coefficient increases very rapidly. This is accompanied by an increase in the density of electrons in the conduction band.

The value of the optical intensity, after passing through a thickness x of material, is given by the relation :

$$I_{(x)} = I_0 \exp(-\alpha x)$$

where $\alpha(\mathrm{cm}^{-1})$ is the absorption coefficient and I_0 is the intensity of the radiation having penetrated into the material (equal to the intensity of the incident wave if one neglects the reflection at the surface).

The penetration of the radiation into the semiconductor can be defined by the penetration length $L = \frac{1}{\alpha}$

In designing a photodetector for a given application, the material to be used must first be determined according to the wavelength of the radiation to be detected. The structure of the photodetector will depend on the nature of the material and in particular on its penetration length.

III. MATERIALS

For the spectral range from 0.4 to 1,5 μ, the material must be a semiconductor whose band gap is on the order of one electron-volt.

Table 1 gives for silicon, germanium, gallium arsenide and indium arsenide, the band-gaps (Eg), the maximum detectable wavelength (λ_{max}), and the refractive index n.

	Si	Ge	GaAs
E_G (300°C) (eV)	1,1	0,65	1,43
λ_0 (μm)	1,2	1,9	0,87
n ($\lambda \gg \lambda_0$)	3,42	4,1	3,40

Table 1 :

Figure 2 gives for silicon, germanium and gallium arsenide the penetration length as a function of wavelength. We note that it would be very difficult to conceive a photodetector if L is greater than several hundred microns or if L is less than 0,5μ . In addition, the photodetector will be faster, if L is small.

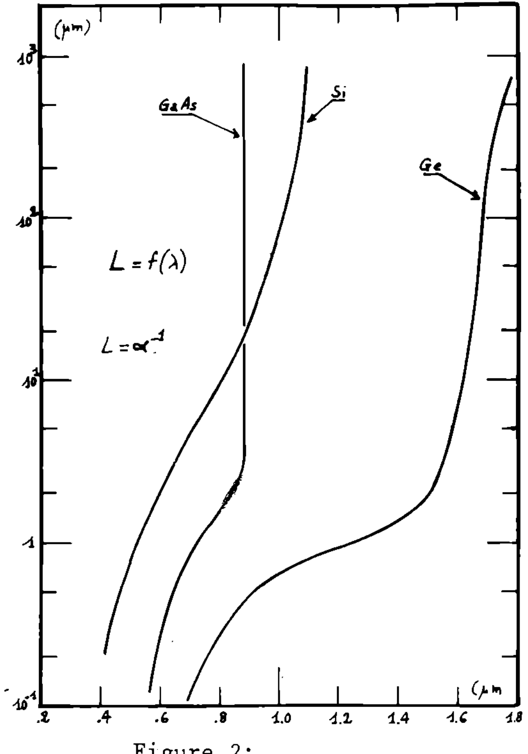


Figure 2:
Absorption length vs wavelength

Other elements must be taken into consideration for the choice of material, the capability of the technology for the fabrication of structures optimized for a given application, the value of the parasitic dark current, which is greater when the band gap is smaller.

▪ IV. STRUCTURES

The absorption of an optical wave by a semiconductor will be accompanied by an increase in the conductivity of the material.

We will only mention photoconductors, which can be used for the simplest photodetecting structures. However these structures do not appear competitive for any application, at least in the spectral range of interest to us and they are almost never used.

Therefore we will discuss only junction photodetectors.

Only a weak leakage current passes through a reverse biased junction. This current is due principally to surface defects and thermal excitation of electrons from the valence band. The absorption of an optical wave adds to the current of minority carriers a current proportional to the number of absorbed photons.

This reverse biased junction has a space charge region, outside of which the electric field is almost zero.

We will distinguish two categories of electron-hole pairs generated by the incident photons :

- Those generated in the space charge region, in which they are accelerated by the electric field ; the resulting response is very fast (on the order of or less than a nanosecond).
- Those generated outside the space charge region which must move by diffusion in the region of zero electric field, before reaching the space-charge region ; the corresponding response time will be on the order of several tens of nanoseconds.

In addition, by applying a sufficiently high reverse bias voltage, the electric field in the space charge region can become large enough that multiplication of the carriers occurs ; the multiplication coefficient being given by the relation :

$$M = \frac{1}{1 - \left(\frac{V}{V_{BR}}\right)^n}$$

where V is a applied voltage , V_{BR} the breakdown voltage and n is the miller coefficient. M can reach a value of several hundred.

The photodetector structure must be optimized to obtain the best compromise for the sensitivity, rapidity, noise. The following factors are to be considered :

- Thickness of the space charge region (compared to the penetration length of the incident radiation).
- Value of the obscurity currents, which will depend on the applied voltage.
- Possibility of obtaining multiplication.

A. Diffused PN junction

Figure 3 shows the simplest structure of a photodetector : the N^+/P junction.

The diffused N ring prevents breakdown at the edge of the N^+/P junction, when high voltage are applied. The breakdown voltage of the N/P junction is higher than that of the N^+/P junction.

The width of the space charge region of the N^+/P junction will depend on the resistivity of the P substrate, and for P substrates having a resistivity above several hundred Ωcm, the space charge region can occupy the entire thickness of the substrate.

However as the resistivity of the substrate increases, the avalanche voltage becomes higher and this is accompanied by an increase in the obscurity current.

In practice, the multiplication phenomenon cannot be utilized for voltages above 200 or 300 V. In this case, the space charge region thicknesses would be on the order of 5 to 15μ for silicon.

Depending on the penetration length of the optical radiation ; one observes a fast response due to absorption in the space charge region, followed by a delayed signal due to the diffusion of

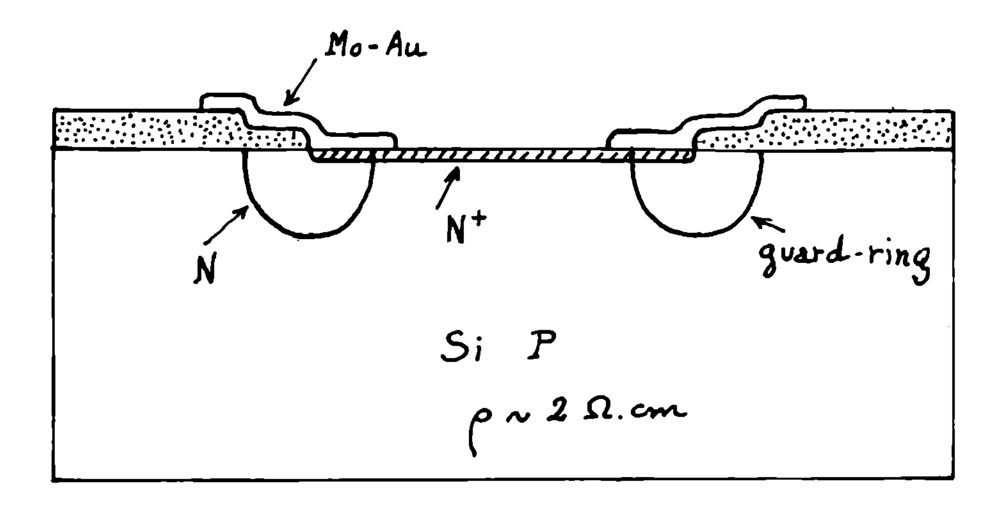


Figure 3 :

N^+P Avalanche Photodiode with diffused guard-ring

carriers generated outside of this zone.

To eliminate this delayed signal, the PN junction can be diffused in an epitaxial layer on a substrate having a very low resistivity. In this case, the lifetime of the carriers in the substrate is very short and the carriers generated in this region recombine before reaching the junction. The response rate is therefore fast, even though the quantum efficiency is low.

B. PIN Junction

The use of weakly doped substrates allows however the extension of the space charge region throughout the full thickness of the substrate. In this case the quantum efficiency is high and the photodetector response is fast, but the multiplication process cannot be used.

C. Schottky junction

In the case of PN or PIN junctions the diffused junctions can be replaced by a metal-semiconductor junction. The operation of a Schottky photodetector will be the same as that explained above.

However the response of the Schottky diodes will be higher for high energy radiation. This kind of radiation is absorbed near the surface of the semiconductor. In the case of diffused junctions, the surface region has a high concentration of impurities and exhibits defects. The lifetime of the carriers in the surface region is therefore short and these carriers will not be collected by the junction.

That is not the case for the Schottky structures.

D. Pπ PN Junction

Figure 4 shows a diode structure designed to obtain the advantages of the avalanche junction and those of a PIN junction. That is, the rapidity and the detectivity.

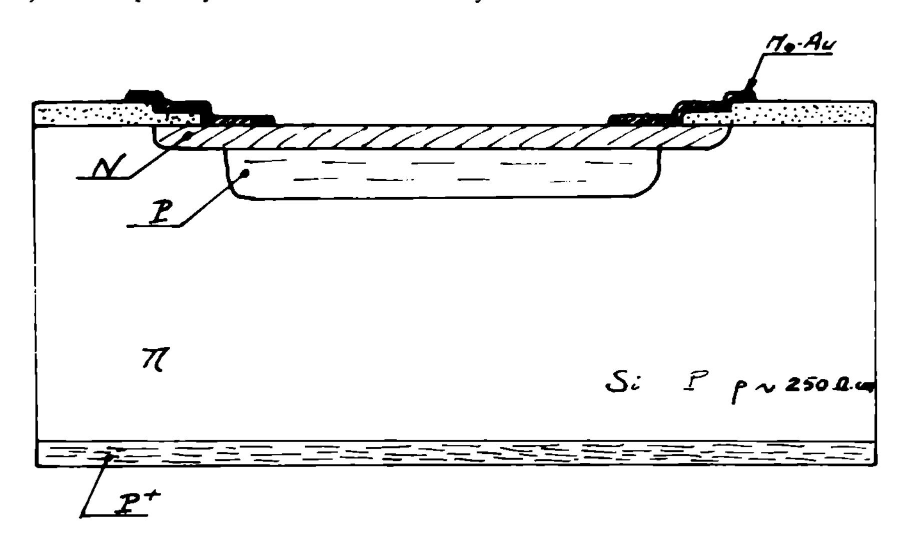


Figure 4 :

PπPN Avalanche Photodiode.

Fabricated using weakly doped π substrates this diode contains the following diffused regions :

P^+ region to facilitate the contact
P region
N^+ region which forms with a P region a PN diode whose breakdown voltage would be on the order of several ten of volts.

Figure 5 illustrates the operation of this device showing the distribution of the electric field E for three values of the applied diode voltage :

$$V < V_{rt} \;,\quad V = V_{rt} \quad \text{and} \quad V > V_{rt}$$

For $V < V_{rt}$: the electric field is confined in the N^+P junction and for a sufficiently high voltage, avalanche occurs.
For $V = V_{rt}$ the space charge region reaches the edge of the P region.
For V slightly larger than V_{rt} : the space charge region occupies the entire region π of the substrate creating an accelerating field for the carriers generated in this zone.

Therefore such a device presents the frequency performance of a PIN diode as well as the multiplication phenomenon.

The difficulty of fabricating such a device lies above all in the adjustment of the doping profiles in the N and P regions so

as to obtain an electric field distribution such as we have just described. Ion implantation seems to be the best method of obtaining a well controlled impurity profile.

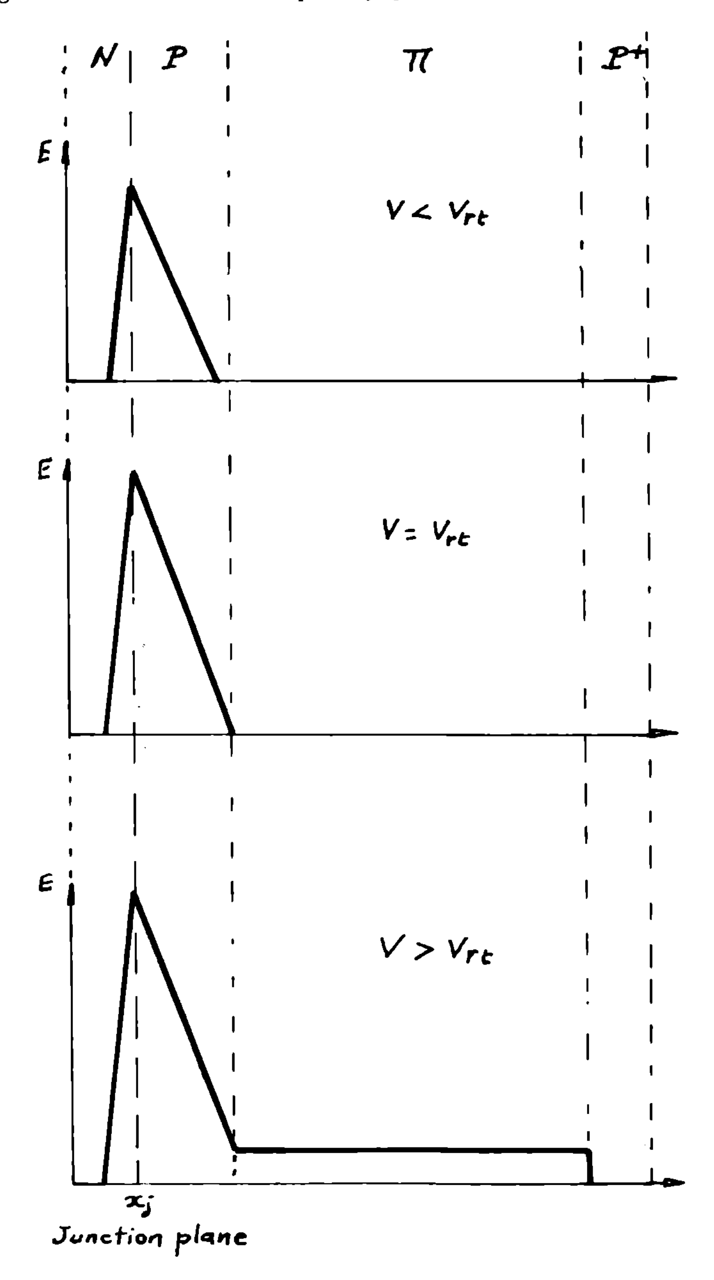


Figure 5 : *Distribution of electric field vs bias voltage V.*

V. TREATMENT OF THE STRUCTURES

The device must be treated in various ways to improve the performance and these treatments will be determined by the utilization conditions.

A. Anti-reflecting layer

Optimization of the index of refraction of the semiconductor material allows a significant improvement (on the order of 25% in the case of silicon) to be gained in the sensitivity of the photodetector. This optimization is obtained through the deposition onto the silicon surface of a thin film of $\lambda/4$ having a bulk index of refraction between that of the semiconductor and that of air. Silicon dioxide can be used, but silicon nitride (Si_3N_4) which has an index of refraction equal to 2± 0,1 , is to be prefered. In addition to its optical role, Si_3N_4 provides a passivation of the diode and improves the reliability.

In the case of a passivation by an antireflecting coating of Si_3N_4, it is not necessary to encapsulate the photodetector, and its performance will not be modified by the presence of a window.

When the spectral range of the detector is wide, the antireflecting layer can be optimized for the region of the spectrum having the lowest sensitivity. This will have the effect of rendering the response as a function of λ constant.

B. Diaphragm

The metalliclayers used for the contacts to the photodetector can also serve as adiaphragm which precisely limits the photosensible zone of the photodetector.

C. Integration

It is often very useful to integrate the photodetector with the amplification electronics and the electronic signal treatment.

This integration is however often impossible in a monolithic form which requires substrate specifications very different from those required for the fabrication of an integrated circuit.

An integration in hybrid form can be obtained in all cases.

VI. PERFORMANCE OBTAINED

We will illustrate the performance obtained with different photodetector structures, presenting the principal characteristics measured for photodetectors available on the market.

Table 6 shows first the characteristics of photodiodes operating in the avalanche modes and fabricated with three types of material : germanium gallium arsenide and gallium indium arsenide.

In this table, for each of the three materials presented, the following factors are given .

- The structure
- The technique of forming the guard ring
- The spectral response
- The diameter of the spectral zone
- The breakdown voltage
- The multiplication coefficient

AVALANCHE PHOTODIODES	Ge	As Ga	$In_x Ga_{1-x} As$
Structure	$N^+ P$	Schottky (Pt, Au)	Schottky (Pt, Au)
Guard ring	diffused	proton bombardment	proton bombardment
spectral range (μm)	0,6 - 1,7	0,3 - 0,9	λ max ⟶ x x = 0,20 0,3 - 1,20
Active Area	ϕ = 250 μm	ϕ = 125 μm	ϕ = 125 μm
Breakdown voltage (V)	50 - 100	50	50
Multiplication factor	50	100	100 - 200
Quantum efficiency	λ = 1,06 μm η = 40% λ = 1,54 μm {η = 60%, f = 0,1 MHz} {η = 25%, f = 10 MHz}	λ = 0,63 μm η = 30%	x = 0,17 / x = 0,2 λ = 1,06 μm: η = 30% / η = 45% λ = 0,9 μm: η = 50% / η = 35%
Sensitivity (A/W)	λ = 1,06 μm σ = 0,35A/W λ = 1,54 μm {σ = 0,75 A/W, f = 0,1 MHz} {σ = 0,30 A/W, f = 10 MHz}	λ = 0,63 μm σ = 0,15 A/W	x = 0,17 / x = 0,20 λ = 1,06 μm: σ = 0,25 / σ = 0,40 λ = 0,9 μm: σ = 0,35 A/W / σ = 0,25 A/W
Response time (ns)	λ = 0,9 μm $\tau_m \leq$ 0,5 ns	λ = 0,68 μm τ = 0,1 ns at M = 100	λ = 1,06 μm $\tau_m \leq$ 0,2 ns for M = 250
Multiplication noise	$P \propto M^3$	$P \propto M^{2,1}$	$P \propto M^{2,3}$
dark current	$i_0 \simeq 5/\mu m$ for V = 15 V		

Figure 6 :

Structure	N^+P	N^+P (epitaxial)	P PN^+	P I N		SCHOTTKY (Si:t)
Guard ring	diffused	diffused	junction N^+	mesa		diffused
Active Area	ϕ = 100 - 750 /um	ϕ = 100 /um	ϕ = 500 /um	ϕ = 3 mm	ϕ = 800 /um	ϕ 200 /um
Spectral range	0,4 à 1,1 /um					0,25 /um à 1,4 /um
Bias voltage Vf(V)	$V \simeq 50$ - 200 V	$V \simeq 50$ - 150 V	$V \simeq 250$ - 400V	$V \simeq 2000$ V	$V \simeq 10$ V	$V \simeq 50$ - 100 V
Multiplication factor	200	200	50 à 100	200		100
Quantum Efficiency (M=1) λ = 0,9um	50 %	16 %	90 %	75 %	70 %	50 %
λ = 1,06um	15 %		30 %	30 %		15 %
Sensitivity (M=1) (A/W) λ = 0,9um	0,35 A/W	0,12 A/W	0,65 A/W	0,55 A	0,5 A/W	0,35 A/W
λ = 1,06um	0,10 A/W		0,25 A/W	0,3 A/W		0,10 A/W
Dark current at V = Vf	10 nA	2 nA	100 nA	700 nA	0,5 nA	100 nA
Response time λ = 0,9 /um	$\tau_m \simeq 100$ ns	$\tau_m \simeq 10$ ns	$\tau_m \simeq 2$ ns	τ_m = 3 ns	$\tau_m < 0,5$ ns	$\tau_m \simeq 100$ ns
Multiplication noise	$P \propto M^{2,5}$	$P \propto M^{2,4}$	$P \propto M^{2,3}$			
Noise equivalent Power	10^{-12} $W.Hz^{\frac{1}{2}}$ at 1 GHz		3.10^{-14} $W.Hz^{-\frac{1}{2}}$ at 1,06 /um (7MHz)	2.10^{-14} W $Hz^{-\frac{1}{2}}$ at λ = 1,06 /um		

Table 7 :
Avalanche silicon detectors.

- The quantum efficiency
- The sensitivity
- The rise time
- The multiplication noise
- The obscurity current

The interest of these photodetectors lies essentially in the rapidity-sensitivity performance.

In particular, one can see that the Ga As Schottky detector has for $\lambda = 0.68\,\mu$, a rise time of 0.1 ns for a total sensitivity of 15 A/W (0.15×100).

For the In GaAs detector, we have a rise time of 0.2 ns for a sensitivity of 100 A/W for $\lambda = 1.06\,\mu$.

Table 7 presents the characteristics of silicon photodetectors operating in the avalanche mode : 6 structures are shown :

- an N^+P junction
- an N^+P junction on an epitaxial substrate
- a $P\pi PN$ structure
- 2 PIN structures
- 1 Schottky structure (or Si Pt)

One can note :

★ the reduction of the sensitivity (from 0.35 to 0.12 A/W) and the increase in speed - (rise time decreasing from 100 to 10 ns) when one passes from the N^+P junction of a homogeneous substrate to the one of an epitaxial layer.

★ the gain in sensitivity (0.65 A/W) and rapidity (2 ns) provided by the structure $P\pi PN$.

★ the rapidity of the PIN not operating in the avalanche mode (0.5 ns), but having a total sensitivity of 0.5 A/W (compared to the total sensitivity of the $P\pi PN$: 50 A/W).

★ the increase in the spectral range for the Schottky diodes.

▪ VII. CONCLUSION

Photodetectors have not yet reached the end of their evolution and better performances can still be obtained. These improvements will be made through the utilization of new materials upon which other techniques can be applied, such as for example the liquid epitaxy of gallium aluminium arsenide. The utilization of heterojunctions between different materials will make possible photodetectors having new properties, such as a selectivity for a given wavelength, or the reverse case, an enlarged spectral range.

LIGHT DEFLECTION BY INDEX CHANGES

J.I. Pankove

R.C.A. Laboratories Princeton

New Jersey 08540 U.S.A.

I. INTRODUCTION

By modulating the refractive index of a transparent material , one can manipulate the propagation of light in a useful way. Light rays can be deflected, their polarization changed, their intensity modulated, and even their frequency can be altered.

The possibility of modulating the amplitude or the frequency of light makes the transmission of radiant energy a useful medium for broad-band communication in either line-of-sight or guided wave transmission. The small size of the wavelength permits a high density of information storage, while the high speed of propagation opens the possibility of rapid information processing. Batches of information, such as pictures can be operated upon either sequentially or simultaneously. A judicious choice of wavelengths allows visible displays for the fastest coupling to the human mind such as at the machine-man interface. Since ultra-short pulses of light can be generated, it is possible to make time-dependent studies with high accuracy. The most precise distance measurements use optical interferometry.

In this paper we shall first examine the various phenomena that can be used to control the refractive index. Then we shall consider several experiments demonstrating the deflection of light. And finally, we shall indulge in speculation about the future of light deflection.

II. FUNDAMENTAL CONSIDERATIONS

The main problem in controlling light is to couple the radiant energy to a small modulator. The radiant energy inside the device is proportional to the volume occupied by the light ; i.e., the diffraction-limited cross section of the light beam integrated over the path length. The efficiency of the device is increased if the light is confined to a smaller volume. Thus, a planar waveguide may reduce the modulation power needed to drive the device by at least one order of magnitude compared to a bulk modulator. Confi-

nement into a one-dimensional channel may further reduce the required driving power by another order of magnitude. A typical state-of-the-art figure of merit of 0.5 mW/MHz is the power needed to provide nearly 100% electro-optic modulation in a 1MHz bandwidth.

The radiation propagates through a solid with a velocity, v, which is related to the velocity in vacuum, c, by $v = c/n$, where n is the real part of a complex refractive index :

$$n_c = n - ik \tag{1}$$

the imaginary part being the extinction coefficient. Losses occur whenever absorption of the radiation is possible : ionization of impurities, excitation of lattice modes, across the gap excitation, scattering at inhomogeneities, resonant transitions and free carrier susceptibility. The latter is avoided by using insulating materials.

The value of the refractive index at some photon energy E can be derived from the knowledge of the absorption spectrum $\alpha(E)$ via the Kramers-Kronig relation :[1]

$$n(E) - 1 = \frac{ch}{2\pi^2} P \int_0^\infty \frac{\alpha(E')}{(E')^2 - E^2} dE' \tag{2}$$

where P is the Cauchy principal value of the integral and the integration is carried over the whole spectrum. A large change in $\alpha(E')$ will affect the refractive index, especially when this change occurs at a value of E' close to E. Note that the denominator makes the contribution of a change in α at E' far from E very small. In other words, the greatest modulation of the refractive index is obtained at photon energies close to those where the absorption coefficient is most strongly affected. In practice, a compromise must be made between a large change in n and a minimum absorption.

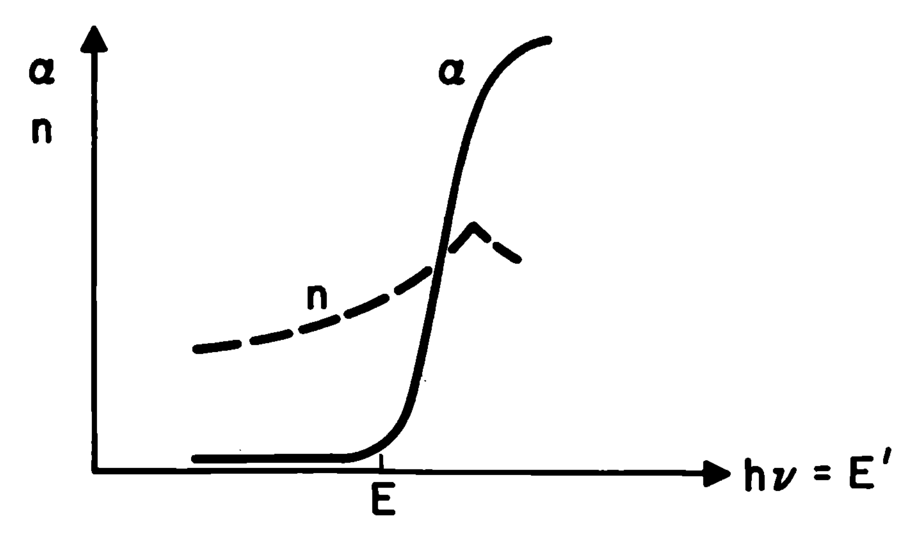


Figure 1 :
Variation of refractive index near an absorption edge.

The parameters which can be used to control the refractive index are the chemical composition of the material (including its doping), free carriers, population inversion, electric and magnetic field, mechanical strains and temperature. We shall now consider these effects individually.

A. Compositional Effects

According to the Kramers-Kronig relation, Eq.(2), the largest variation of the refraction index is obtained near the absorption edge of the material. Figure 1 illustrates a typical interdependence between n(E) and α(E). The location of the absorption edge in the energy spectrum can be adjusted by selecting the composition of the crystal. This can be readily achieved in some semiconductor alloys such as $Al_{1-x}Ga_xAs$ where the absorption edge moves to higher energies as the Al concentration increases, causing a drop in the refractive index at photon energies where the GaAs is still transparent (Fig. 2). This compositional control is used to make waveguides that confine the radiation in injection lasers. An extension of this technique could lead to curved waveguides that can deflect light in any arbitrary direction [2]. However, these are static devices which, once constructed, do not modulate the deflection.

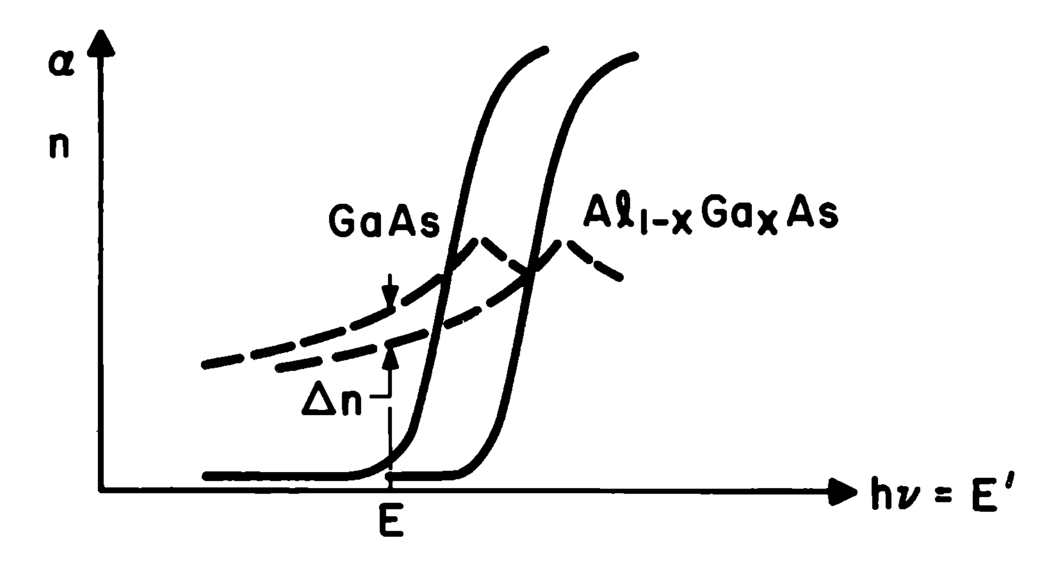


Figure 2 :
Effects of composition on absorption and on refractive index. This figure illustrates why a GaAs layer sandwiched between two $Al_{1-x}Ga_xAs$ layers forms a waveguide.

B. Free-Carrier Effects

Free carriers produce two effects which cooperate in reducing the refractive index in the useful region of the spectrum. One of

these effects is free-carrier absorption which increases as some power, p, of the wavelength (λ^p), where $1.5 < p < 3.5$ [3]. This is shown in Fig.3. In practice, one would work at a photon energy, E, where the absorption is minimal. Most of the free-carrier absorption

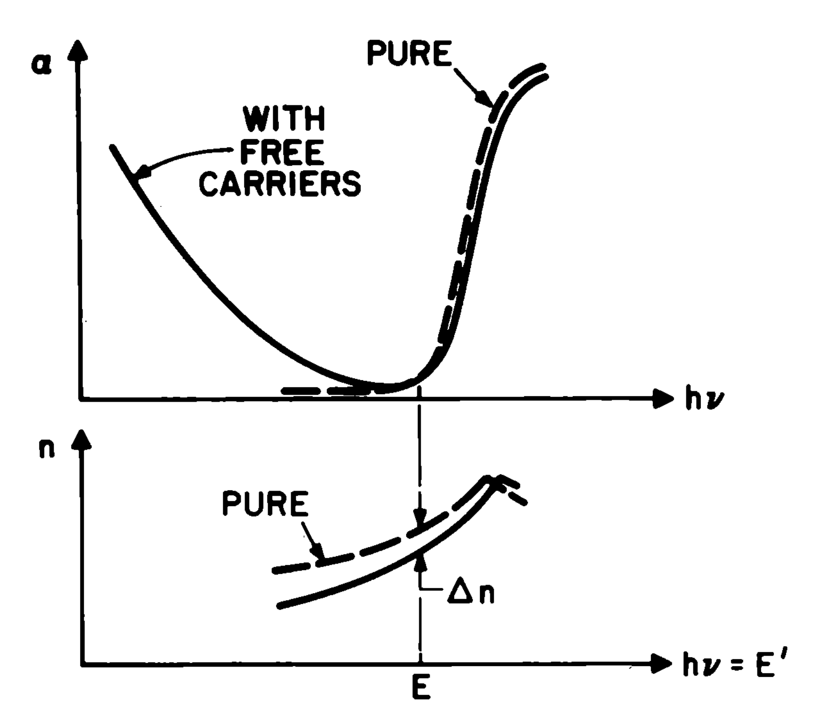


Figure 3 :
Effect of free carriers on the absorption spectrum and on the refractive index.

occurs at E' < E ; therefore, the free-carrier contribution to the Kramers-Kronig integral is negative, causing a decrease in refractive index. The second effect is the band filling effect or Burstein-Moss shift [4,5] which makes states near the bandedge unavailable and thereby shifts the absorption edge to higher energies. A drop in $\alpha(E')$ at E'> E is also a negative contribution to the refractive index at E.

C. Impurity Effects

A high concentration of impurities in a semiconductor has a complex effect on the absorption spectrum and therefore on the refractive index. A single type of impurity, either donor or acceptor, contributes free carriers and thereby depresses the refractive index by free-carrier absorption and by the Burstein-Moss shift. A further effect is "band-tailing" which results from the perturbation of the crystal by local stresses and by Coulomb interactions. [6] This effect spreads the absorption edge over a larger portion of the spectrum than in the case of a pure material (Fig. 4). The band-tailing effect is most pronounced in compensated materials where a

high concentration of both donors and acceptors is present without contributing free carriers. [7] Band-tailing makes a positive contribution to the refractive index in the spectral range of interest (see Fig.4). The combination of free-carrier absorption, band-filling, and band-tailing can either lower or raise the refractive index depending on which of these effects is dominant. Figure 5 shows the energy dependence of the refractive index in n-type GaAs as measured by Sell et al. [8] Note that at 5.9×10^{17} electrons/cm^3 band-tailing dominates, whereas at 6.7×10^{18} electrons/cm^3 free carriers and band-filling effects take over. In p-type material, where free-carrier absorption is very strong, the refractive index is depressed.

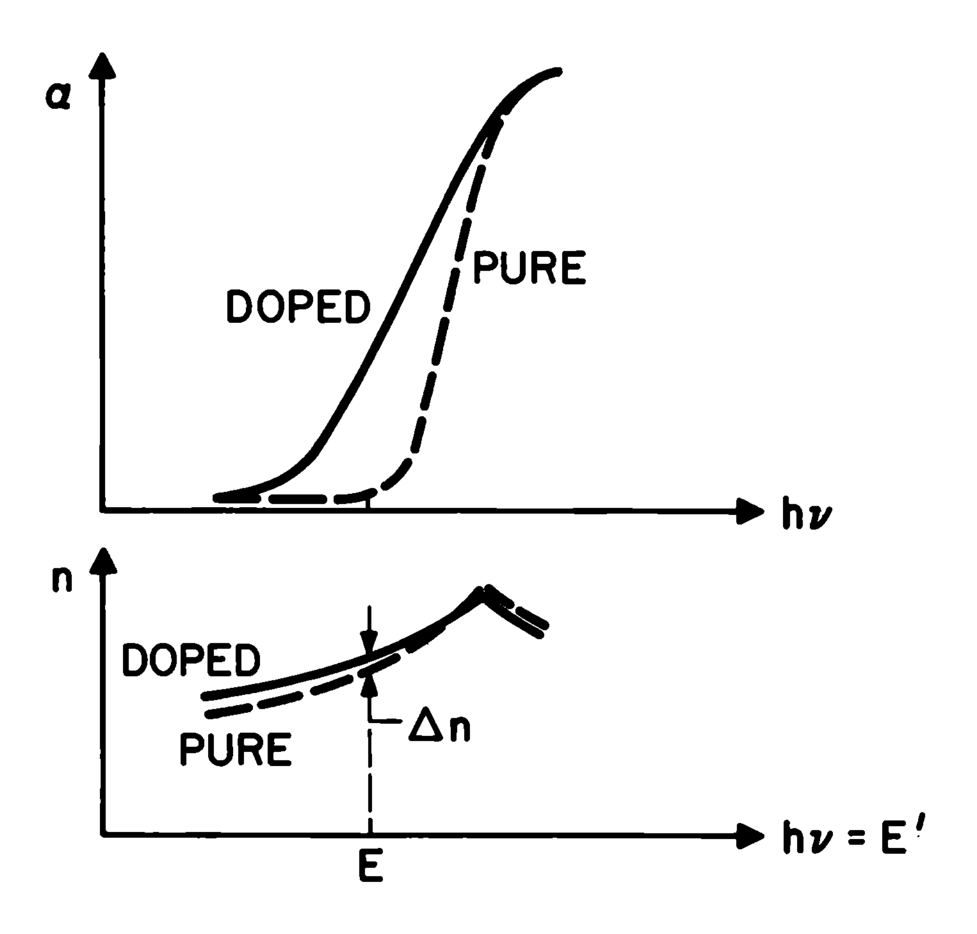


Figure 4 :
Effect of doping on the absorption edge (low energy tailing) and on the refractive index.

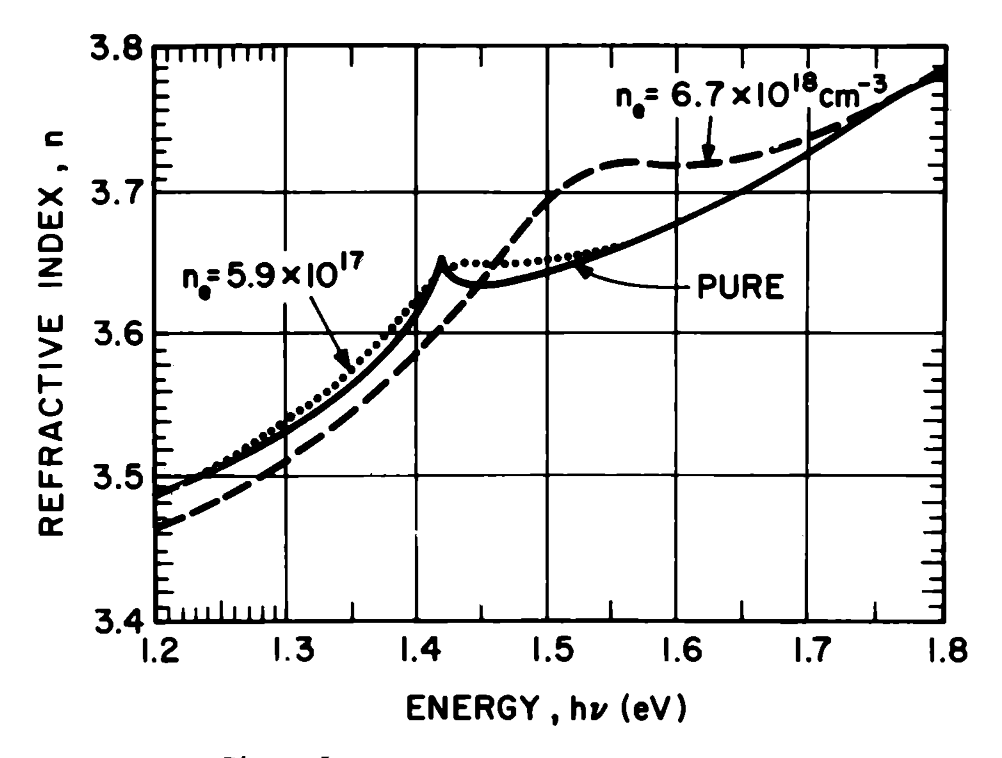


Figure 5 :
Refractive index of pure and doped GaAs. Room temperature data of Sell et al.[8]

In general, impurities, like composition, are fixed constituants of the material and therefor^ are not modulated. However, some impurities such as lithium c… migrate through a crystal under the influence of an electric field.[9] This impurity drift mechanism could provide a means for changing reversibly, but slowly, the refractive property of a crystal.

D. Double Injection Effects

When both electrons and holes are simultaneously present in a large concentration ($\gtrsim 10^{18} cm^{-3}$) in a semiconductor, in addition to the above free-carrier absorption and band-filling effects, population inversion is obtained. This condition of population inversion can be achieved by electrical injection, electron bombardment, or optical pumping. Population inversion creates a spectral region where the absorption coefficient is negative (Fig. 6). There, the refractive index is increased- an effect which partly accounts for light trapping in homogeneous semiconductor lasers. Note that the radiation pattern from the early injection lasers, from electron-beam pumped or from optically-pumped semiconductor lasers is similar and corresponds to a diffraction-limited radiation pattern emerging from the end of a $\sim 1\ \mu m$-wide waveguide. The p-n junction and the pumped surface of the semiconductor are obvious discontinuities forming one boundary of the waveguide. The

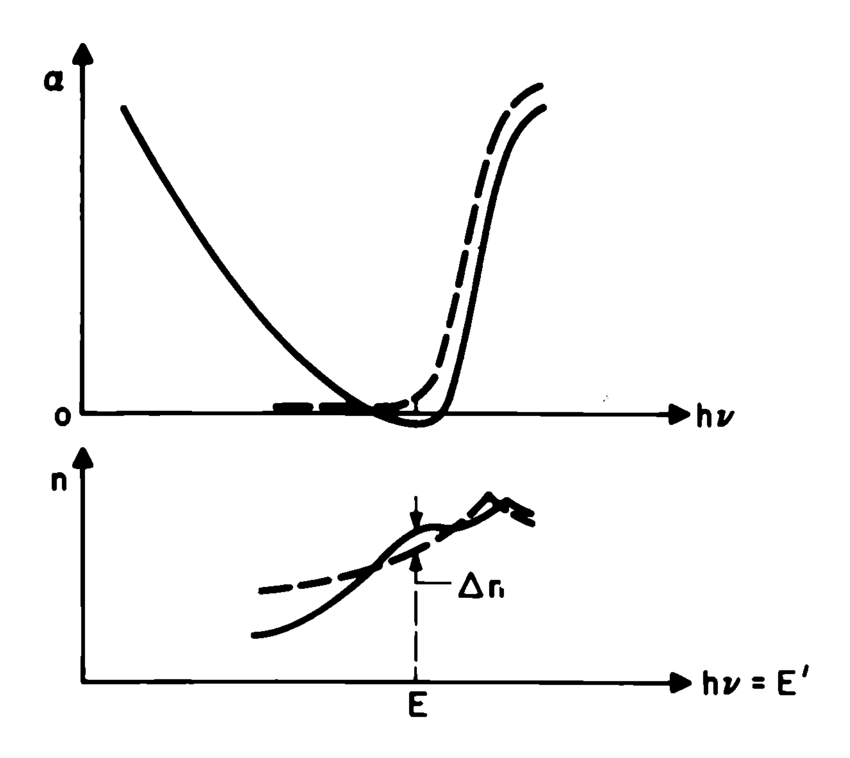


Figure 6 :
A negative absorption coefficient due to population inversion by intense double injection increases the refractive index.

other change in refractive index 1 μm away may be accounted for by the drop in population inversion about one diffusion length away from where the high density of pairs occurs. In the case of electron-beam pumped lasers, the population inversion can be extended to greater depths by increasing the beam voltage. In that case, the radiation pattern can be narrowed considerably.[10]

E. Thermal Effects

The energy gap and therefore the absorption edge of semiconductors is temperature dependent, some with a positive temperature coefficient (e.g., HgTe for which $dE_g/dT = 5.6 \times 10^{-4}$eV/°K), but most with a negative temperature coefficient (e.g, Se for which $dE_g/dT = -1.4 \times 10^{-3}$eV/°K). As a consequence of the Kramers-Kronig relation, the refractive index is also temperature dependent. Light can be deflected by controlling the temperature gradient in the crystal. [11]

Nonradiative recombination is a fast process which generates phonons ; i.e., heat. Nash [12] has calculated that a radiative efficiency of 50% in a GaAs injection laser causes a 1.5°K rise in temperature inside the active region. This localized heating corresponds to a change in refractive index $\Delta n/n$ of 2.4×10^{-4},

which is comparable to the increase in refractive index due to population inversion. [12]

F. Strain Effects

Hydrostatic pressure increases the density of the compressed medium and causes an increase of the refractive index. This is the elasto-optic effect. In semiconductors, hydrostatic pressure can either expand or contract the energy gap. For example, in GaAs $dE_g/dP = 11 \times 10^{-6}$eV/bar. [13] Near the absorption edge, the refractive index will experience a corresponding change : decreasing with a positive pressure coefficient, increasing for a negative dE_g/dP. Thus, Stern finds that for GaAs, near the absorption edge, $dn/dP \approx -9 \times 10^{-6}$/bar. [14] As we shall see below, the strain dependence of the refractive index, the "acousto-optic" effect, may be the most useful means for deflecting light.

In anisotropic or birefringent materials, more complex interactions are possible in which the polarization of the light is affected. Unaxial stress also produces a complex interaction because both compression and dilation occur simultaneously in different directions. Hence, some states may move to higher energies while other states move to lower energies, warping the spatial configuration of the refractive index at a given photon energy.

G. Field Effects

A strong electric field shifts the absorption edge to lower energies and thereby increases the refractive index. This field-induced shift of the absorption edge is the Franz-Keldysh effect ; [15,16] it has been observed in Si, [17] GaAs, [18] and Ge. [19] A built-in field exists at a p-n junction due to the locally steep impurity gradient. (This field exists even when no voltage is applied across the junction.) A change in refractive index of 0.02 has been reported in Ge due to the built-in field. [20] A gradient of electric field (and therefore of refractive index) can be produced by a judicious design of electrode geometry (e.g., a point contact). A light scanner with a resolution of 1000 using the Franz-Keldysh effect was proposed by Boer. [21]

In materials which are piezoelectric, part of the index change may result from the induced strain and other effects rather than from the Franz-Keldysh effect.

Many materials exhibit the Pöckels effect which is electric field induced birefringence and is also called the linear electro-optic effect. In isotropic crystals, when the electric field is zero, the index of refraction is isotropic and the dielectric ellipsoid is a sphere. When an electric field is applied along a selected direction which we shall call z, the index of refraction n_z does not change ; but in the transverse plane the refractive index is altered by an amount Δn, proportional to the electric field ξ_z. The modification of the dielectric sphere occurs in such a way that the refractive index changes by $\pm\Delta n$ in mutually perpendicular directions. For example, $n_x = n + \Delta n$ and $n_y = n - \Delta n$.

The actual orientation of the crystal for a maximum effect is guided by a choice of crystallographic directions. The optimum choice depends on the crystal-group symmetry to which the material belongs. The magnitude of the linear electro-optic effect is :

$$\Delta n = \frac{n^3}{2} r\xi_z \qquad (3)$$

where r is the electro-optic coefficient. Table I lists the electro-optic properties of several semiconductors.

TABLE I

ELECTRO-OPTIC PROPERTIES OF SEMICONDUCTORS

Material	*Electro-optic coefficient at room temperature r(cm/V×IO^{-10})*	*Refractive index at zero field n*	*n^3r (cm/V×IO^{-10})*	*References*
GaAs	1.6	3.34	59	22
ZnTe	1.4	2.79	34	22
ZnS	2.0	2.37	27	23
CdTe	6.8	2.6	120	24
CdS	5.5	2.3	66	22

There is also a second order electro-optic effect, the Kerr effect, which varies quadratically with the field. In some materials, the selffocusing of intense laser beams can be explained by the Kerr effect : [25,27] The high electric field of a coherent wave can increase the refractive index in the direction of light propagation by the transverse electro-optic effect in materials having a large Kerr coefficient. A beam with a gaussian intensity profile generates a maximum increase in refractive index at the center of the beam where, therefore, the phase velocity is decreased. As the beam propagates, the equiphase surface becomes more concave towards the direction of propagation. Hence, the rays converge toward the region of highest intensity.

Magnetic fields can also induce a refraction anisotropy. The Faraday effect is linear with the magnetic field and affects light propagation in the direction of the magnetic field, while the Voigt effect is the magneto-optic equivalent of the Kerr effect. Unfortunately, at present, magneto-optic modulators compete with electro-optic modulators only in the infrared.

III. EXPERIMENTAL LIGHT DEFLECTORS

A. Electro-optic Deflection

Beam deflection can be obtained with the electro-optic modulation of prisms. A large number of these can be stacked and cascaded to provide a large deflection and a great number of resolvable spots. [28] Unfortunately, of all the possible schemes for beam deflection, varying the refraction of a prism requires, by far, the most driving power.

One can take advantage of double refraction to make a digitized

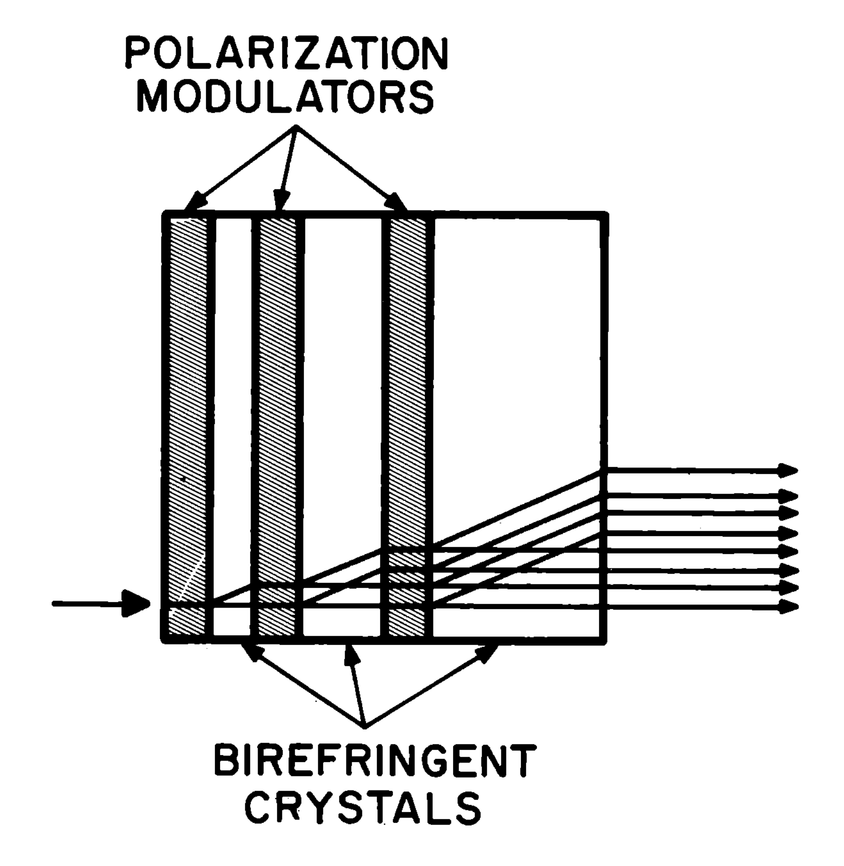


Figure 7 :
Digital birefringent deflector, the polarization modulator selects whether the ordinary or extraordinary ray is to be transmitted. Only one of the eight beams shown is allowed to go through the stack.

beam deflector. A light beam incident on a birefringent material at an angle to the optic axis of the crystal (it must be neither parallel nor perpendicular to the optic axis) breaks up into two beams which travel at an angle to each other and emerge from two different spots. The separation of these spots is proportional to the path length inside the birefringent crystal. The ordinary ray is plane polarized perpendicularly to the optic axis while the extraordinary ray is plane polarized perpendicular to the ordinary ray. The two rays emerge displaced with respect to each other and with their plane of polarization mutually orthogonal but travelling in the same direction, parallel to the incident beam. An electro-optic modulator in front of the birefringent crystal can select which of the two beams will emerge. To get additional displacements, one can stack m birefringent crystals each preceded (or followed) by an electro-optic polarization selector. This will produce 2^m spots (a set of 7 will allow 128 spots). Note that, as shown in Fig 7, successive birefringent crystals must be made twice as long as the preceding one to accumulate the displacements. The most readily available birefringent material is calcite. Because of materials costs and because of the large power needed to drive the electro-optic modulator, this method does not appear practical.

As mentioned in the introduction, the most efficient operation is obtainable with a waveguiding modulator. In the device of Fig. 8, an array of interdigitated electrodes is placed on one surface of the waveguide. [29] When a voltage is applied to this array, the electric field along the array alternatingly changes polarity. Hence, the refractive index fluctuates as a function of position by $\Delta n = \pm \frac{1}{2} rn^3\xi$ about the zero field value n. If the waveguide is thin compared to the interelectrode spacing, s, the refractive index is practically uniform in the volume between the electrodes. Such an alternation of high and low refractive indices forms a grating which can scatter light like the periodic planes of a crystal can scatter X-rays. There is a Bragg condition :

$$2 \sin\theta = \lambda /2s \qquad (4)$$

where θ is the angle of incidence with respect to a constant index plane. The length, L, of the iso-index planes must be large enough to make a thick grating which scatters light into a single first order diffraction ; i.e., $2\pi \lambda L >> (2s)^2$.

Laser light is coupled into the waveguide by a prism coupler (at the left side of Fig. 8) and is coupled out of the waveguide with another prism coupler (at the right side of Fig. 8). Inside the waveguide, the incident beam is then deflected by an angle 2θ when a voltage is applied across the array. Note that since the deflection is determined by the spacing, s, of the electrodes, a fixed angle of deflection is obtained. This applied voltage determines only how much of the incident radiation is coupled into the scattered beam. The intensity of the deflected beam is given by :

$$I = I_0 \sin^2 aV \qquad (5)$$

where I_0 is the maximum intensity of the undeflected beam and a is a constant. If a detector is located so as to collect either the transmitted or the scattered beam, this device is mostly an intensity modulator capable of 100 % efficiency. [29]

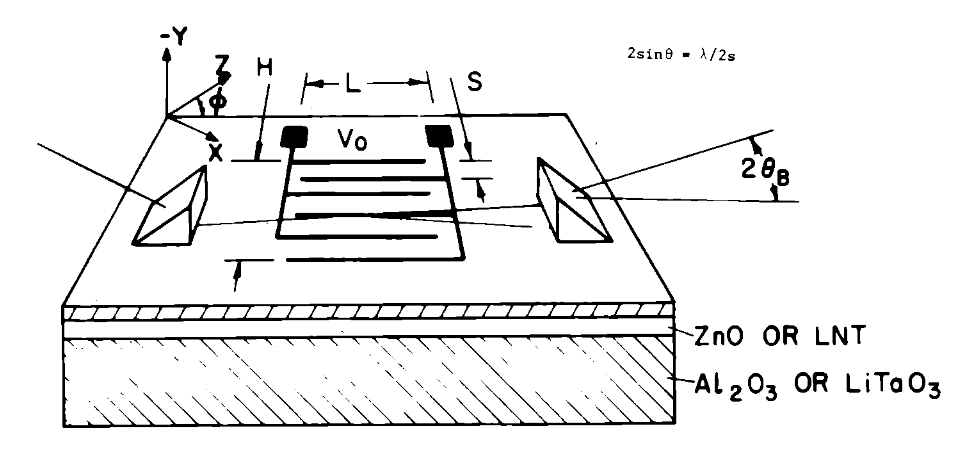


Figure 8 :
Schematic diagram of electro-optic waveguide grating modulator. (Courtesy J.M. Hammer).

Figure 9 shows the percentage of the light scattered into the first order beam as a function of the applied voltage obtained with the most recent improvement in material technology by Hammer and Phillips. [29] A number of such two-position deflectors can be cascaded to provide a plurality of deflection angles. To satisfy the Bragg condition, successive identical electrode arrays must be tilted by $2m\theta$, where m is the sequential order of the array.

Electro-optic deflectors and modulators are base-band devices which can be driven at high frequency. An instrument-limited response time of 3 nsec has been measured, and operation well in the GHz range is expected. [30] The upper frequency limit is determined by the allowable power dissipation in the driving circuit which delivers a power $P = \frac{1}{2} CV^2\Delta f$, where C is the device capacitance and Δf the bandwidth. With the most advanced material, lithium-niobate-tantalate (LNT) the driving power requirement was less than 0.2 mW/MHz. [29] Electro-optic semiconductors tend to be lossy and therefore may have a lower frequency limit than electro-optic insulators such as LNT. The latter effect, losses due to residual free carriers in semiconductors, can be utilized as a further means of intensity modulation of the deflected beam. [31] The free carriers can be generated by UV irradiation. This technique is of particular interest to photonics.

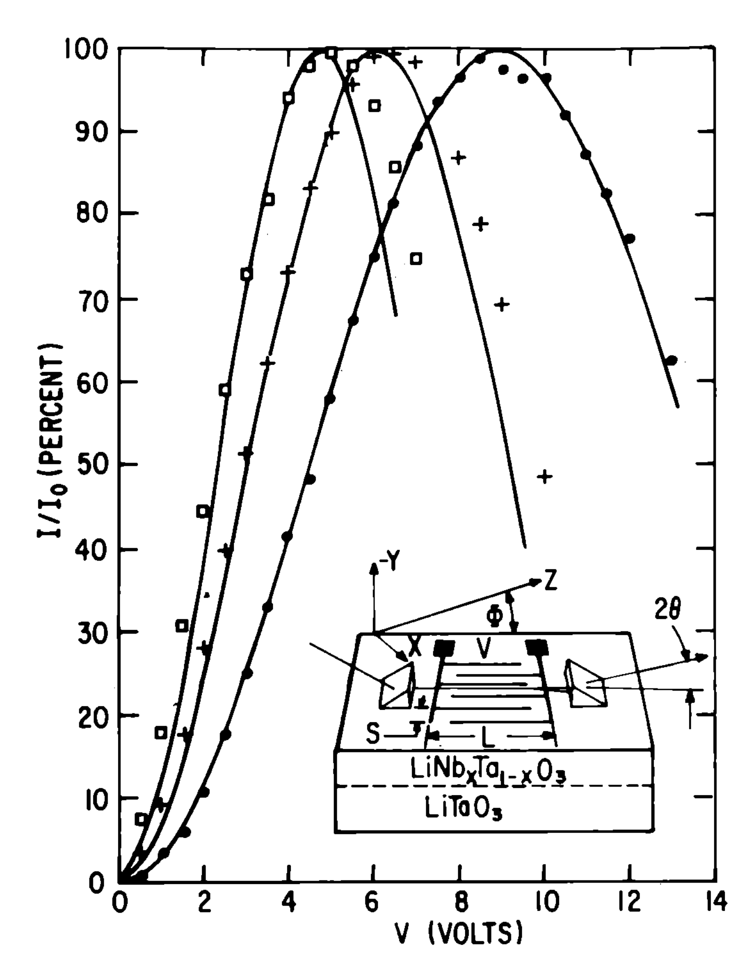


Figure 9 :
Percentage of light diffracted as a function of voltage. Guided light is diffracted through an angle 2θ when a voltage is applied to the inter-digital electrodes. S is 7.6 μm and L is 0.3cm. Open squares 4976 Å (He-Se laser), crosses 5598 Å (He-Se laser), and solid circles 6328Å (He-Ne laser). The solid curves are plots of $sin^2(aV)$ normalized to the data at 75% (Courtesy J.M. Hammer).

B. Acousto-optic Deflection

A continuous deflection of the light beam can be obtained with an acousto-optic device. The scattering of light by soundwaves was predicted by Brillouin [32] in 1922 and demonstrated ten years later. [33,34] For our purpose, a transducer generates planar soundwaves which propagate in the transparent medium of the modulator. The soundwaves form alternating regions of high and low pressure which

modulate the refractive index (Fig. 10).

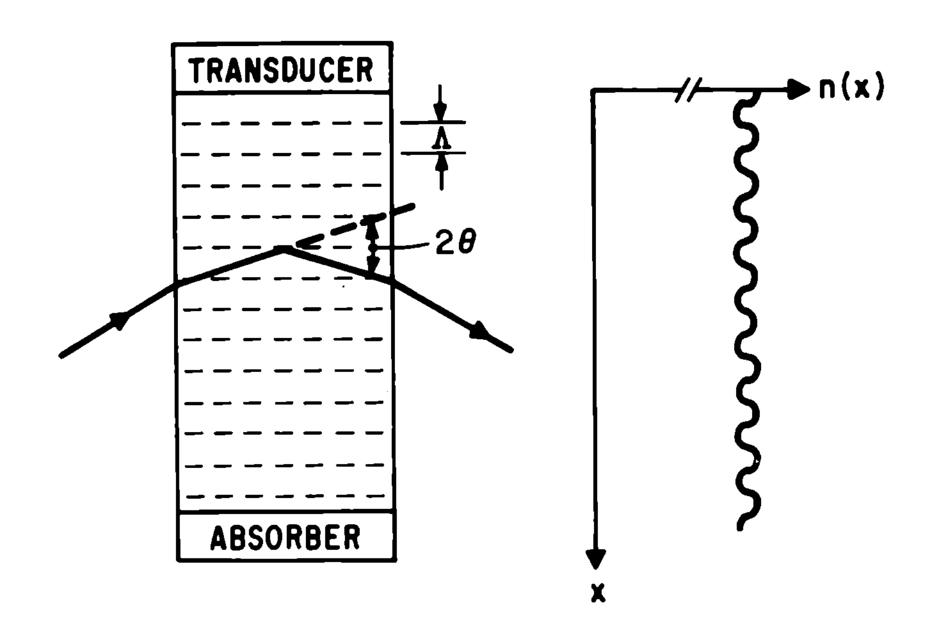


Figure 10 :
Schematic diagram of acousto-optic Bragg deflector and distribution of refractive index along the device.

In general, regions of high hydrostatic pressure have a higher density and a higher refractive index. However, in semi-conductors, where the acousto-optic effect might modulate light at photon energies close to that of the energy gap, the change in refractive index may be more sensitive to the pressure dependence of the gap than to the elasto-optic or density effect. Thus, in semiconductors having a positive pressure dependence of the energy gap, the refractive index may drop with increasing pressure.

In either case, soundwaves generate a refraction grating which can deflect light by an angle 2θ according to the Bragg condition :

$$2 \sin \theta = \lambda / \Lambda \quad (5)$$

where Λ is the acoustic wavelength.

The deflection is proportional to the frequency, f, of the soundwave since $\Lambda = v/f$ where v is the sound velocity. The angle θ is small therefore $\sin \theta \sim \theta$ and the deflection is :

$$2\theta = \lambda f/v \quad (7)$$

This relation suggests that to get a large deflection, one must find a material with a slow sound propagation velocity. However, there

are other requirements, such as low acoustic losses. In an exhaustive study of materials suitable for acousto-optic modulators, Pinnow [35] concluded that the most useful materials would probably be those having a high density, such as those containing lead cations. However, some trade off may be required between efficiency and maximum usable frequency. At the end of 1969, it appeared that only the longitudinal acoustic wave would generate a useful change in refractive index. Shear wave would not affect the material's density, and therefore any change in refractive index would be a second order effect. However, it was soon found that paratellurite (TeO_2) is an excellent acousto-optic deflector. [36] TeO_2 has a slow shear wave which allows the device to be small (for use with small optical apertures) and an anisotropic index splitting which allows operation in the most useful frequency range. The shear wave converts the ordinary ray into an extraordinary ray which, as we shall see, is crucial in making the acousto-optic deflector a practical device. As shown in Figure 11, the first order scattering,

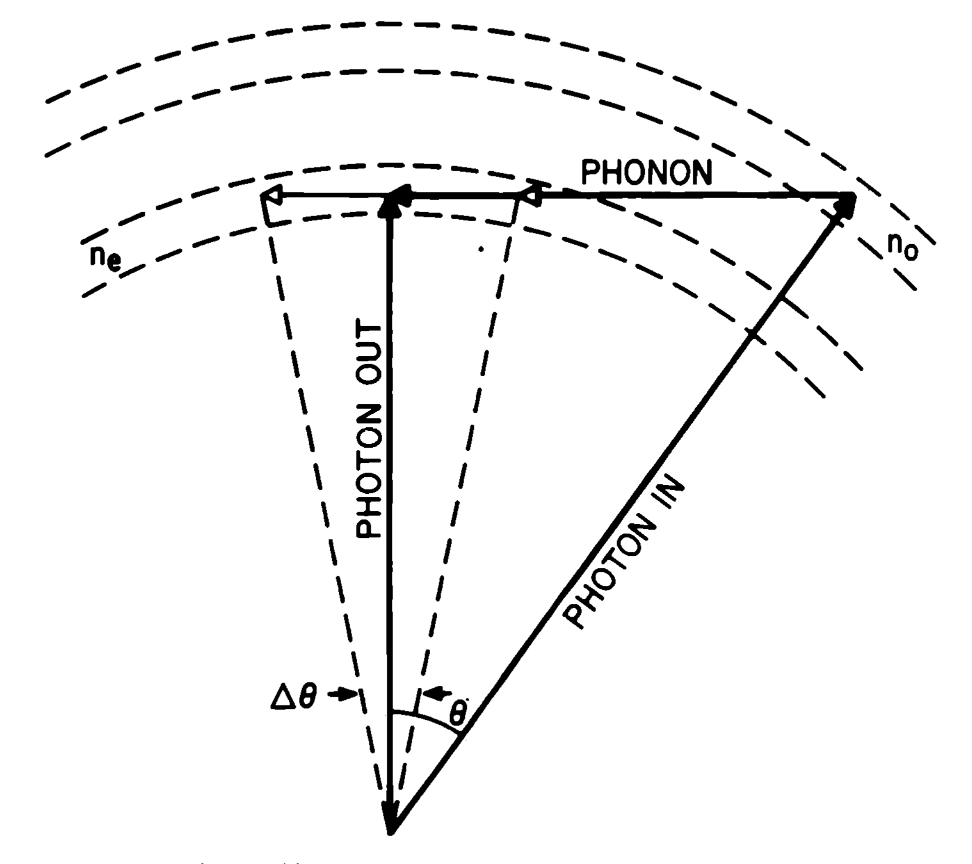


Figure 11 :
Momentum conservation in photon-phonon interaction with transverse phonon (shear wave) in birefringent crystal.

a one-phonon process conserves momentum and changes the type of ray. The separation of unwanted orders is readily accomplished by the use of an appropriate aperture. Since the Bragg condition is satisfied only for one angle θ, the intensity of the deflected beam should drop off rapidly on either side of the Bragg satisfying deflection. This limitation can be eased to some extent either by the diffraction spread of the soundwave direction due to the finite size of the transducer (or by focusing the soudwave) or by focusing the incident beam inside the deflector using a short focal length lens to increase the range of incident angles. [37] This approach is penalized by lower efficiency. Another solution for increasing the efficiency of the deflector over a larger angular range is to use a large transducer and to tilt the plane of the acoustic wave in proportion to the sound frequency. The steering of the soundwave can be achieved with a phased array of transducers.[38] In anisotropic acousto-optic materials, the closure of the momentum vectors occurs as a tangent on a sphere (Fig. 11), which allows some lattitude in the phonon wavelength. The actual spread of usable deflection (the roll-off characteristic) is further broadened by an uncertainty in the value of the refractive index which depends on the intensity of the soundwave (a varying quantity).[39] This fuzziness of the index is shown in Fig. 11 by the thickness of the two index spheres. Hence, a larger range of deflection is possible (broadened by $\Delta\theta$) as indicated by the dashed arrows.

The measured roll-off characteristic for TeO_2 is shown in Fig.12. The dip in the middle of this characteristic is due to the onset of a two-phonon interaction which couples energy to the second order diffraction mode. The polarization of the scattered ray does not change since the second phonon swings the ray by 2θ back onto the n_0 -index sphere.

The acousto-optic deflector is a base-band device : The deflection can be maintained in a steady direction with a CW acoustic drive or it can be varied at some high frequency. The shortest response time is determined by the time required to fill the device with a soundwave of the desired frequency ; i.e., t = (width of device)/v. Practical devices operate with an access time of less than one microsecond. For projection TV application, the shortest deflection time needed is the scan retrace which lasts 12 μsec, well within the capability of present state-of-the-art deflectors.

For projection TV applications, the most important consideration is the number of spots that can be resolved per maximum deflection. If the width of the beam is D, its diffraction limited spread is λ/D. The number N, of resolvable spots is then the total deflection divided by the beam spread :

$$N = \frac{\lambda\left(\frac{1}{\Lambda_1} - \frac{1}{\Lambda_2}\right)}{\lambda/D}$$

where Λ_1 and Λ_2 are the extreme values of the sound wavelengths. Therefore :

$$N = \frac{D}{v}\left(\frac{v}{\Lambda_1} - \frac{v}{\Lambda_2}\right)$$

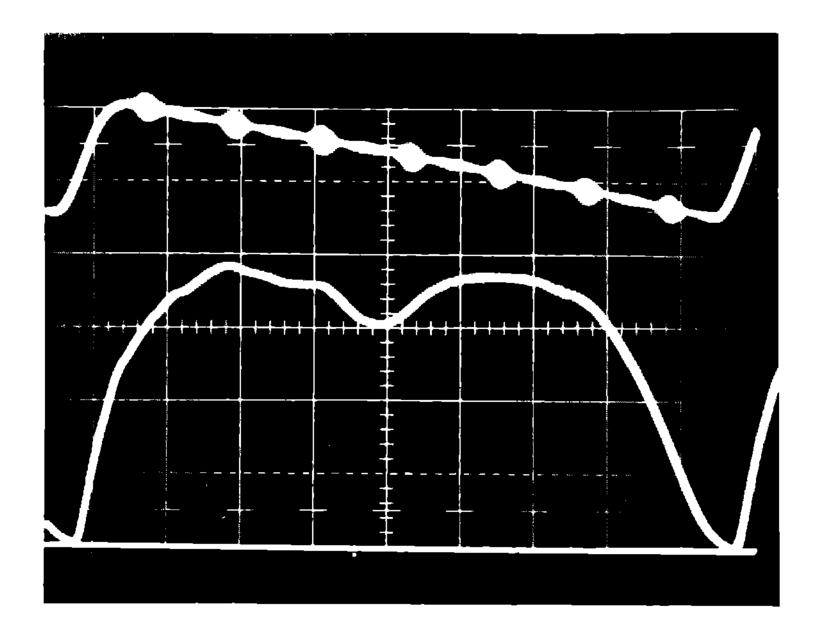


Figure 12 :
Roll-off characteristics of the TeO_2 horizontal deflector. Horizontal scale is approximately 7 μsec, or approximately 8.75 MHz acoustic sweep per major division. Upper trace is horizontal deflection ramp signal with rf frequency markers spaced 10MHz apart superimposed on it ; extreme left marker : 90MHz, extreme right marker :30MHz. Lower trace is deflected light intensity when the average deflection efficiency is 50%. Optical wavelength is 4880Å.(Courtesy I. Gorog) [39]

Note that $D/v = \tau$ is the transit time of the soundwave across the beam. Hence, the number of resolvable spots is $\tau\Delta f$ or, in other words, the change in number of diffraction elements intercepted by the light beam. A resolution of 600 spots has been achieved using a 6 mm diameter light beam and a TeO_2 deflector driven at 60±30MHz which provided a 2.2° internal deflection. [39]

A consequence of the interaction of optical and acoustical waves is the resulting change in the frequency of the scattered light. Energy conservation requires that, depending on the relative direction of propagation of the two waves, sum or difference frequencies are obtained. If light and sound have components propagating in the same direction, the difference frequency is obtained, whereas if they tend to propagate in opposite directions, the sum frequency is obtained. This frequency shift permits an easy noise-free detection of the modulation. [40] For example, mixing a deflected portion of a coherent light with the undeflected remainder of a laser beam at the photodetector reproduces the modulating signal.

Note that the electro-optic modulator discussed previously has a structure which could generate acoustic waves. One might worry that the acousto-optic effect might interfere with the electro-optic effect. In general, the acousto-optic effect is much weaker than the electro-optic effect, except when the crystal is driven to an acoustically resonant frequency.

Strong acoustic waves can be generated in those acousto-optic materials which are piezoelectric. Thus, in the acousto-electric effect [41] the conversion of electrical into acoustical energy results from a cooperative interaction between electrons and phonons. When electrons are accelerated to a velocity exceeding the sound velocity, they transfer kinetic energy to the lattice in the form of phonons. During this buildup of phonons, the electron velocity saturates (decreasing mobility or increasing resistivity). With a constant voltage applied, a high field acousto-electric domain develops which then travels along the specimem with the velocity of sound. Within the domain, many of the effects mentioned in Section II cooperate in increasing the refractive index : increased absorption by phonon-induced tailing of the absorption edge, thermal shift of the absorption edge (phonons are synonymous with heat), piezo-optic stress, and Franz-Keldysh effect. [42] The material becomes birefringent in the domain, hence its motion can be followed by strobing the polarized visible light (Fig 13).

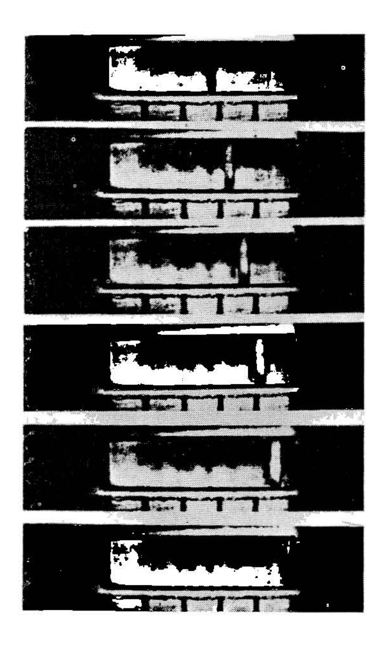


Figure 13 :
Motion of birefringent acousto-electric domain. The voltage was applied two microseconds before the first frame. Interval between pictures is 0.25 μsec. Scale below crystal is 1 mm/div.[43]

The motion of the acousto-electric domain can be used to make a scanning light source. Successive elements of an array of p-n junctions along a semiconductor bar can be caused to emit light during the passage of the scanning acousto-electric domain.[44,45,46]

Inside a transparent liquid, the interference pattern between the soundwave transmitted through (or reflected by) a specimen immersed in the liquid and the unscattered soundwave forms a volume hologram. Under some conditions, such a hologram can be read with a coherent optical beam to form a visible image that can be examined in real time. [47] The ripples that are generated at the gas-liquid interface ; e.g., the free surface of a water cell, can be read in real time as a two-dimensional hologram. [48]

C. Magneto-optic Deflection [49]

Ferromagnetic films consist of domains in the shape of parallel stripes having identical widths and extending the length of the film (Fig. 14). Adjacent domains are magnetized in opposite directions. These domains exhibit a high Faraday rotation. If F is the

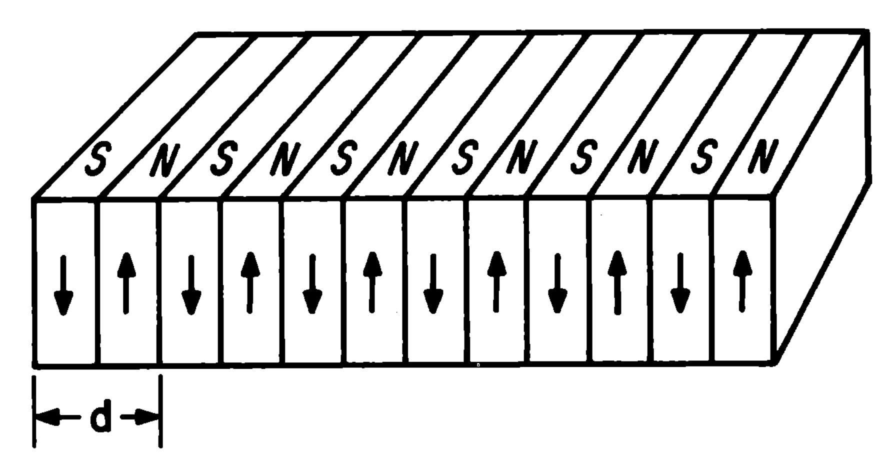


Figure 14 :
One dimensional ferromagnetic stripe domain grating.

Faraday constant and t the film thickness, the angle of polarization of the light is rotated by $\phi = Ft$ in one set of domains and by $-\phi$ in the adjacent set of domains. Because of the differential rotation of adjacent stripes, the film acts as a grating which can deflect light by an angle $2\theta_m$ such that $\sin\theta_m = m\lambda/2d$ where m refers to the mth order diffraction and d is the domain spacing or periodicity. The spacing of the stripes can be varied by applying a magnetic field in the plane of the film, thus varying the angle of deflection. A variation of spacing by more than one order of magnitude has been observed. [49] The deflection is restricted to an angular spread which minimizes the contribution of higher

order modes, hence to a range of angles of about $\lambda/2d_0$, where d_0 is the spacing at zero field.

Since d_0 is of the order of several microns, with a 1cm diameter beam it is possible to resolve several hundred spots.

With present state-of-the-art, magneto-optic deflectors are inefficient in the visible (< 0.1%) but could be practical in the infrared.

D. Double Injection Beam Scanner

It is possible to displace the lasing beam inside the laser cavity by controlling the spatial distribution of the population inversion in the cavity. As we saw in Section II-D, the refractive index is increased wherever the absorption coefficient becomes negative. This principle was successfully demonstrated by Fenner using an injection laser. [50] The structure , shown in Fig. 15, consists

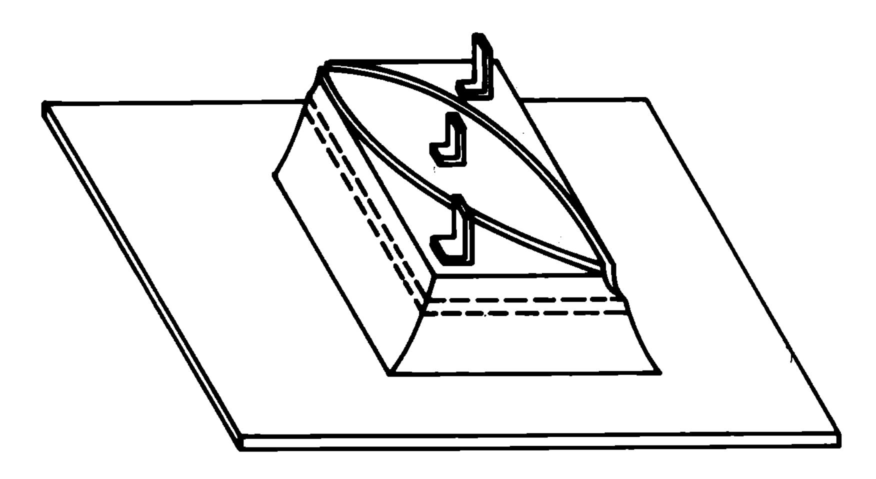


Figure 15 :
Structure of beam scanning injection laser. (Courtesy G.E. Fenner) [50]

of a GaAs laser in which the current distribution is made nonuniform by sectioning the device into three optically coupled but separately excited regions. The separation is achieved by etching grooves into the p-type region. Changing the relative intensities of the currents injected into the different sections of the junction causes the region of maximum gain to move across the width of the junction. The criterium for lasing is given by :

$$R \exp \int_0^{\ell} G(x)\, dx = 1 \qquad (8)$$

where R is the facet reflectivity, G(x) is the local net power gain per unit length and ℓ is the length of the cavity. In the case

illustrated, this condition can be rewritten :

$$\sum_i G_i \ell_i = \ell n \frac{1}{R} = f(y) \qquad (9)$$

If the lengths ℓ_1 and ℓ_3 of the end sections have a quadratic dependence on the position y and if $2G_2 > G_1 + G_3$, it can be shown that f(y) in Eq. (9) is maximum at :

$$y = \frac{W}{2}\left[1 + \frac{G_1 - G_3}{2G_2 - (G_1+G_3)}\right] \qquad (10)$$

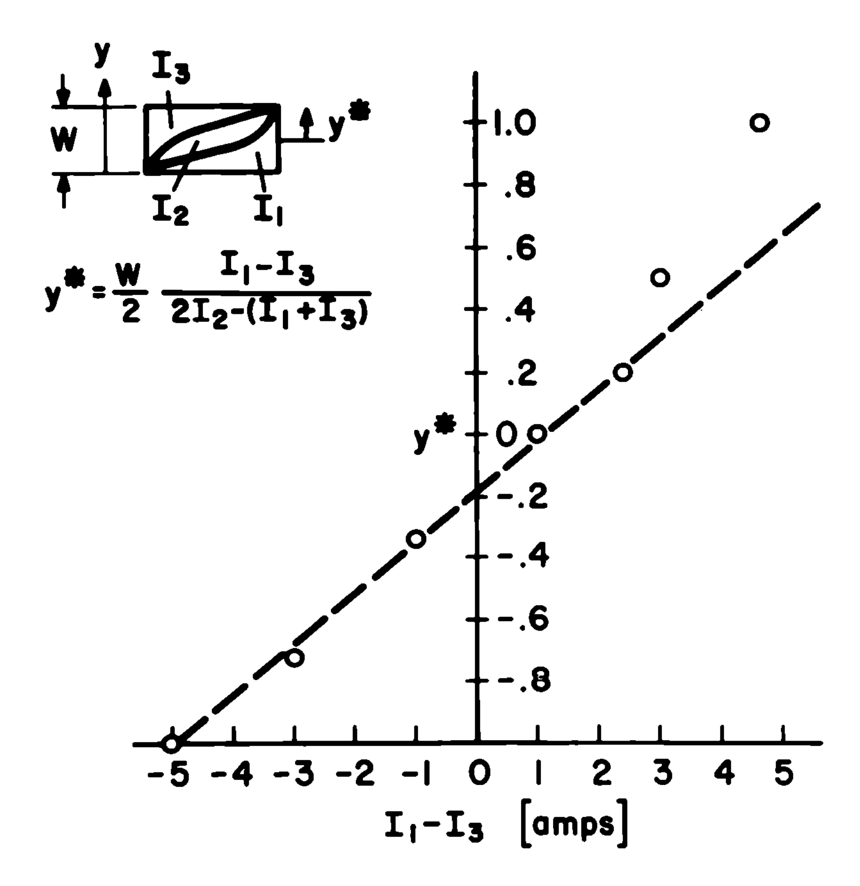


Figure 16 :
Displacement of lasing filament as a function of differential injection pumping of end sections. The quantities involved in the plot are identified on the insert.

Clearly, by manipulating G_1 and G_3, one can move the laser beam on either side of the mid-point at w/2. Note that near lasing threshold

$$G \approx \beta j - c$$

where β is a constant, j is the current density and α is the effective loss per unit length. Therefore, to a first order approximation, the G_i's in (10) can be replaced by the current in the corresponding section. Then, the shift of the beam is accomplished by adjusting the currents in the end sections.

Figure 16 shows the results obtained by Fenner. The ordinate is marked in units of w/2, while the abscissa represents the difference, $I_1 - I_3$, of the measured currents in the end sections. The

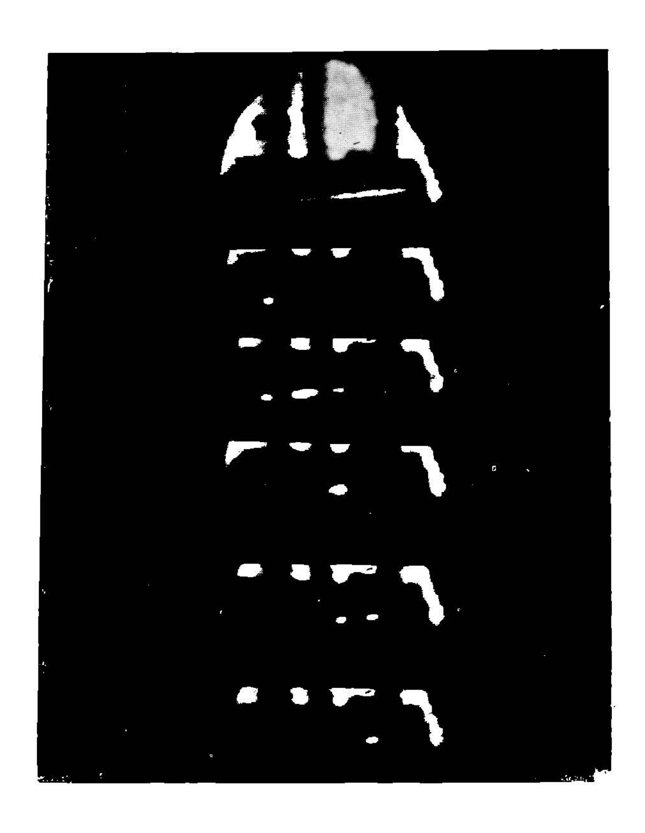


Figure 17 :
Multiple exposure of laser facet under different drive conditions. The top photograph shows the spontaneous emission along the p-n junction below lasing threshold. The other photographs show the lasing filament or filaments moved to various positions.

displacement is linear with $I_1 - I_3$ over a range which is remarkably large considering the state of injection laser technology at that time. Figure 17 is a multiple exposure photograph of the laser diode facet under different bias conditions. The top photograph shows the diode operated in the spontaneous mode with nearly uniform emission over the width of the junction. The other exposures show the displacement of the lasing filament as $I_1 - I_3$ is varied.

E. Ray Deflection by Free-Carrier Injection

We have seen in Section II-B that in the spectral range below the energy gap, the presence of free carriers causes a drop in refractive index. Consider the structure of Fig. 18 in which a kerf separates the electrical contacts to the two p- and n-regions which remain optically and electrically coupled. When the current flows from p_1 to n_1 or n_2, carrier injection occurs at the p-n junction with a distribution that maximizes at the junction and decay exponentially away from the junction. Therefore free-carrier absorption and band filling produce a gradient of refractive index which can be distorted by varying the relative proportion of current flowing to n_1 or n_2 . A preliminary experiment was made with the structure of Fig. 18 without using an external light source ; the light was generated mostly at the p_1n_1 junction. Two

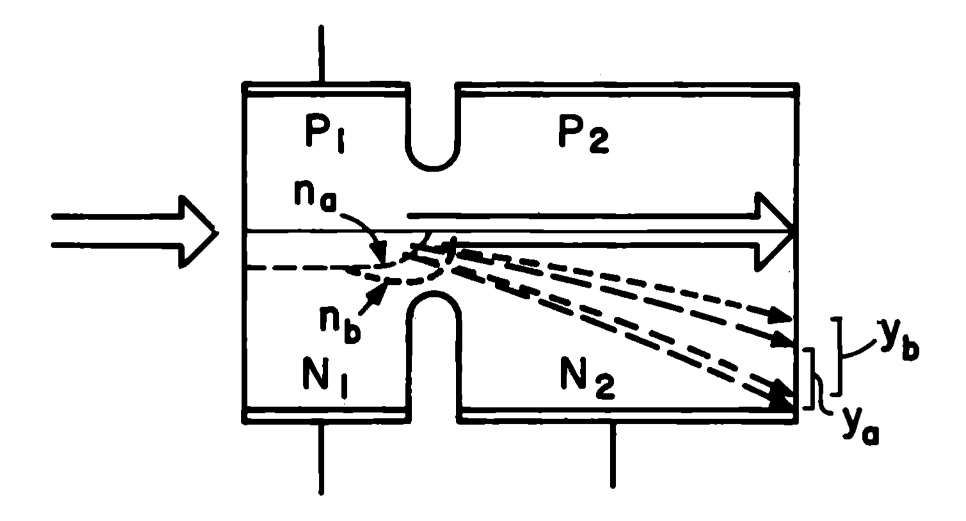


Figure 18 :
Structure for light deflection by free-carrier injection. Part of the incident light follows the p-n junction, the rest is deflected into regions y_a *and* y_b *by iso-index surfaces* n_a *and* n_b*, respectively.* n_a *results from current* $p_I \to n_I$ *, while* n_b *results from some of the current being drawn to* n_2 *.*

bands of light were obtained (Fig. 19) : one at the junction, and the other (more diffuse) near the contact end of the n-type layer. The currents through n_1 and n_2 were in the form of pulses, the coincidence of which could be controlled. In this way, the total

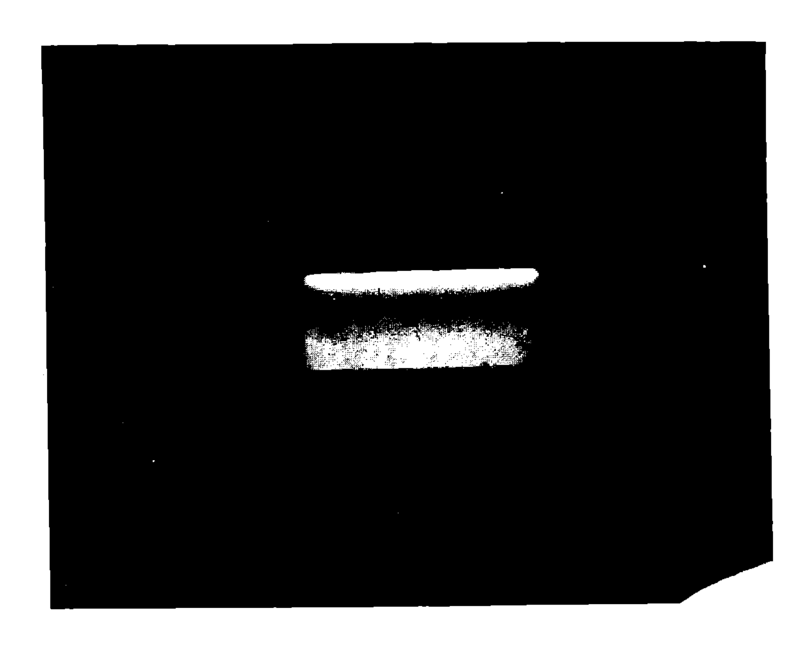


Figure 19 :
I.R image of right-hand facet of the device of Fig.18. The p-type region above the bright band at the p-n junction is not visible.

amount of light generated was constant, but during the coincidence of the pulses, the flow pattern was distorted. A slight deflection was observed as a narrowing of the dark band. From the motion of the lower luminous edge and taking into account the sample geometry, a deflection angle of 1° was estimated.

Further experiments with a more sophisticated design should be performed to verify this preliminary observation and to enhance the effect. For example, the arrangement of Fig. 20 should be tried using a p-n junction in Si-doped GaAs where the index discontinuity can be minimized (Si is an amphoteric impurity which can be dominantly either an acceptor or a donor depending on the temperature during the growth of the crystal). [51] Here, the curvature of the iso-index profile can be inverted when the current is changed from $p_1 \rightarrow n_2$ to $p_2 \rightarrow n_1$.

In another approach, the free carriers can be introduced in a waveguide by optical excitation with light of greater than bandgap energy. The light to be deflected would have a lower than bandgap energy and would be transmitted along the waveguide. Depending on

the geometry of the excited area, the light can be refracted, deflected, focussed or defocussed as shown in Fig 21. Similar results would be obtained with electrical injection ; however, optical injection has greater appeal to an all optical "photonic" process.

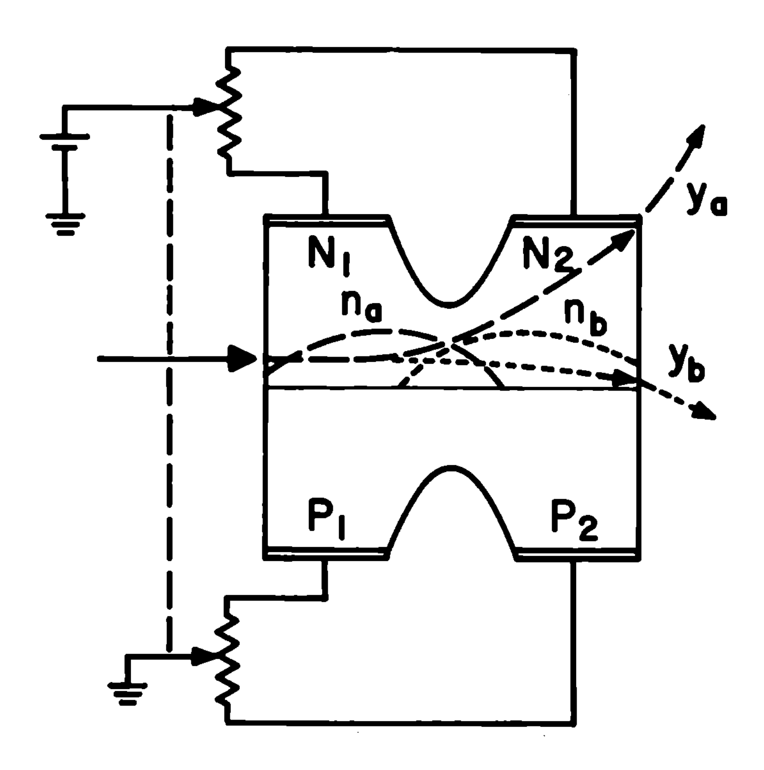


Figure 20 :
Proposed structure for a light deflector by free-carried injection. The current flows to an adjustable degree from p_I to n_2 or p_2 to n_I defining iso-index surfaces ranging from n_a to n_b, respectively.

The most convincing demonstration of beam deflection by optical generation of free-carriers is the experiment of McFee, et al.[52] Their experimental arrangement is shown in Fig. 22. The 10.6 μm output of a CO_2 laser is coupled to a GaAs waveguide in the vicinity of an optically pumped region 2 mm long. The 10.6 μm radiation propagates along the edge of the optically pumped region. At a pump power of 5×10^4 W/cm^2 ; i.e., at about 5×10^{17}carriers/cm^3 , a 3% drop in refractive index is expected at 10.6 μm in GaAs. This localized decrease in the refractive index deflects the 10.6 μm beam away from the pumped stripe. An external deflection of 1° has been obtained (see Fig. 23). This deflection is very fast (< 6 n sec). Because of free-carrier absorption, the deflected beam is attenuated.

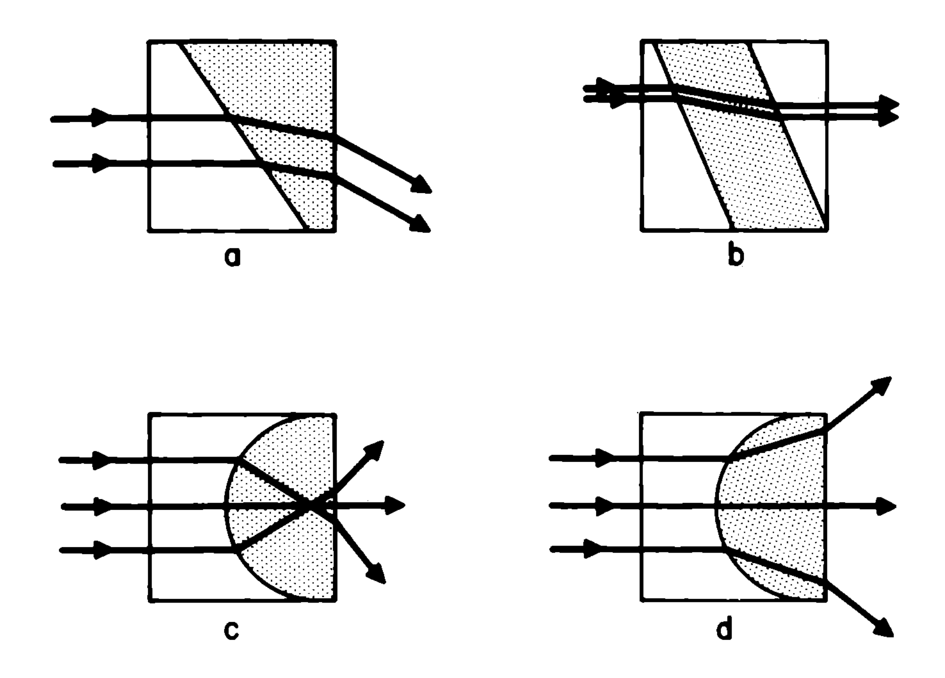


Figure 21 :
Optically controlled ray deflection. The shaded area is irradiated with greater than bandgap light to generate free carriers.(a) beam deflection, (b) beam displacement, (c) focussing and (d) defocussing.

An intense light beam at photon energies near the absorption edge of a semiconductor will create many free carriers and associated effects, such as bandfilling and trapping, which may result in induced transparency (the absorption edge shifts to higher photon energies). The exciting light will then experience a lowering of the refractive index. Ashkin, et al. [53] have studied the defocussing of a laser beam in CdS and have found that the refractive index can drop by as much as one part in a thousand. This effect was produced in the following way. A laser beam at near-gap energy was focussed on a platelet of CdS (Fig. 24). The profile of the beam was examined at some distance from the crystal as the intensity of the laser beam was varied. As shown in Figs.25 and 26, the beam appeared to expand when its intensity was increased, suggesting that the refractive index at the center of the focussed spot was decreasing so that the illuminated region became a divergent lens. Eventually (at 400 mW) the crystal became transparent and the lens effect disappeared. The induced transparency might be expected from bandfilling . However, this effect lasted

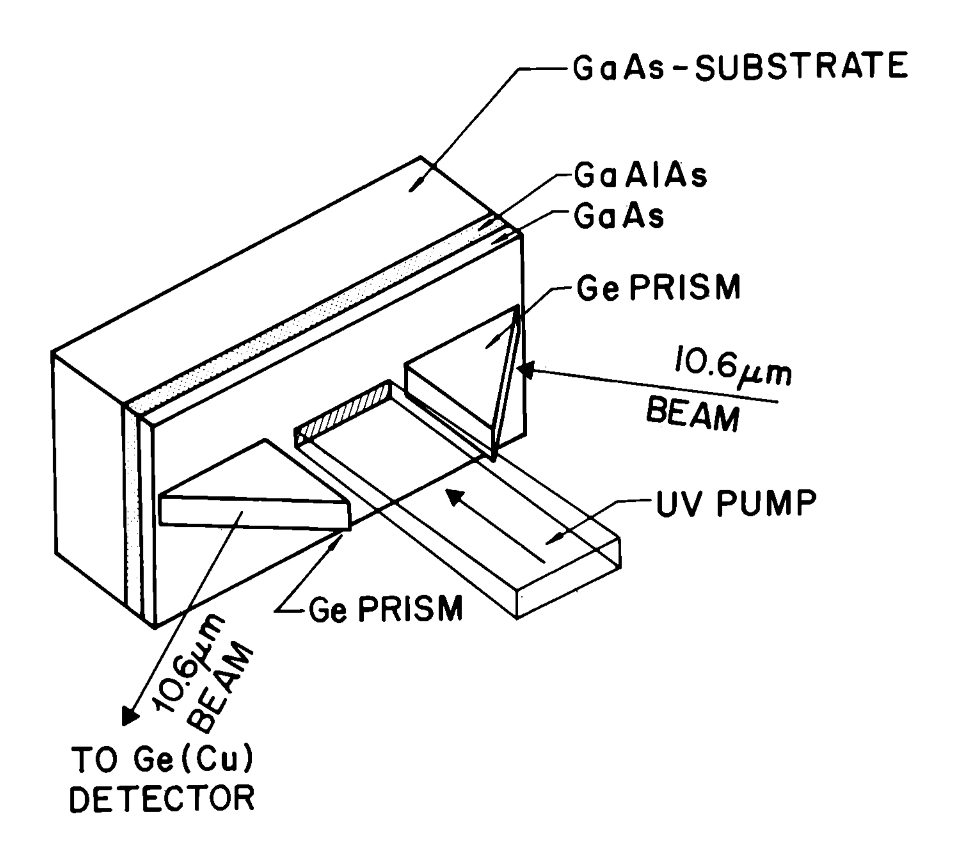


Figure 22 :
Experimental configuration for modulation of a IO.6 -μm guided wave by optical injection of free carriers in a GaAs-AlGaAs heterostructure waveguide. The uv pump beam illuminates the GaAs waveguide surface in the region between the prism couplers. (Courtesy J.H. Mc Fee) [52]

for several seconds after excitation, indicating that carrier trapping in impurities may be involved. Although the neutralization of self-compensating impurities, with a consequent elimination of band tailing appears to be the operating mechanism, more work is needed to fully understand this effect.

The opposite effect or self-focussing of an intense laser beam has also been reported ; [25,27] but, as we saw at the end of Section II-G, self-focussing can be due to the transverse electro-optic effect.

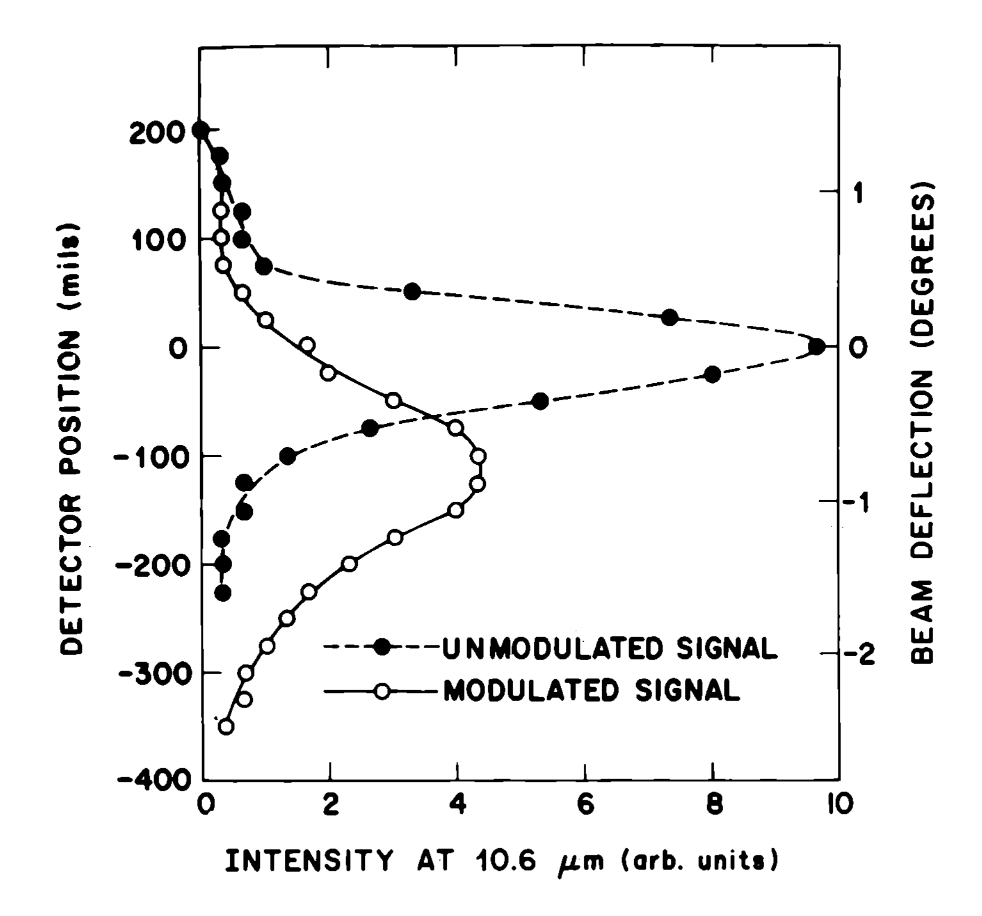


Figure 23 :
Spatial profiles of the output 10.6 -μm wave demonstrating deflection due to optical injection of free carriers. The deflection is measured by moving the apertured Ge(Cu) detector across the 10.6 -μm beam parallel to the plane of the GaAs film waveguide. Angular resolution is 0.45° . (Courtesy J.H. Mc Fee). [52]

F. Thermally Induced Deflection

This effect has been demonstrated in GaAs by Liu and Walters. [11] The current through a slab of GaAs caused heating at the center of the crystal. The contacts at the ends of the slab formed two symmetrical heat sinks. A temperature gradient was thus generated. The light beam from a 1.15 μm HeNe laser traversing the crystal in a region of high thermal gradient was deflected as shown in Fig. 27.

Thermally induced increase in refractive index has been found to be responsible for the self-focussing of laser beams in certain glasses. [54] .

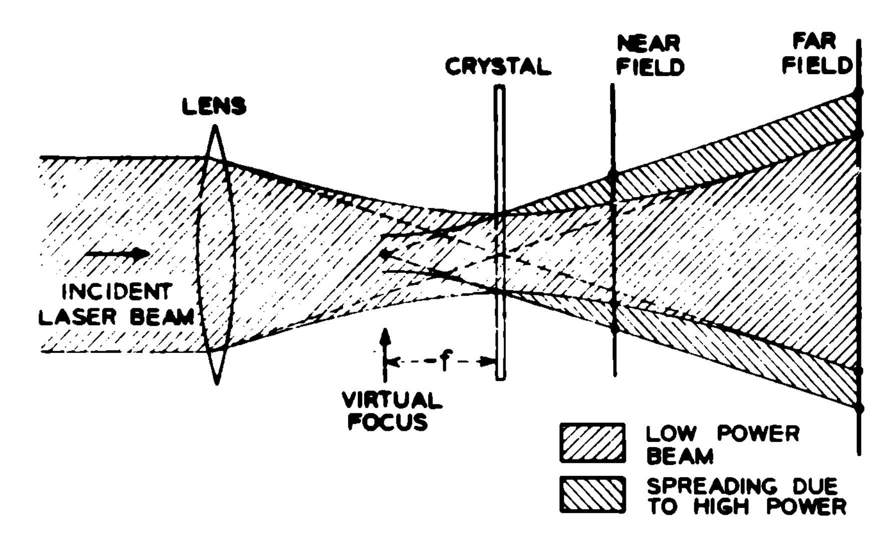


Figure 24 :
Schematic description of experiment. The near field of the incident laser beam is the roughly parallel region at the focus over which the beam radius does not vary by more than $\sqrt{2}$. The far field is the region of uniform angular divergence. The effect of increased laser power on the beam sizes in the near and far fields is sketched. (Courtesy A. Ashkin) [53]

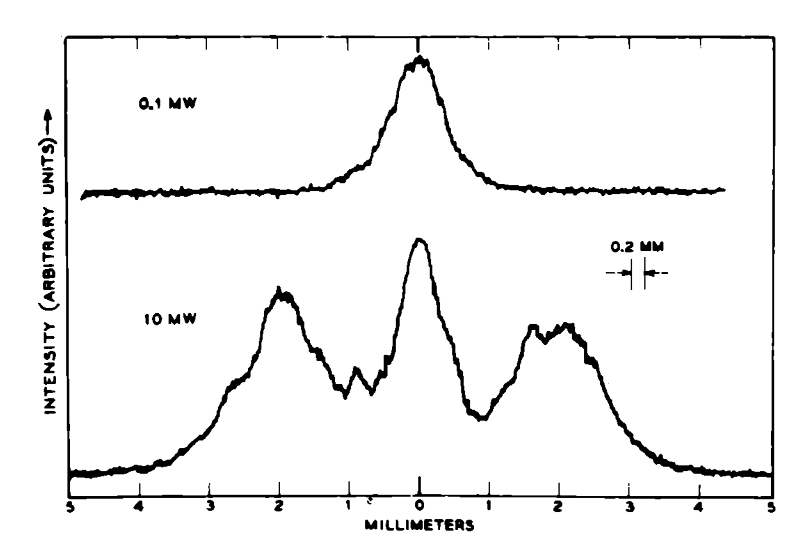


Figure 25 :
Scan of far-field image at 0.I mW and IOmW of incident power using slit width indicated. (Courtesy A. Ashkin). [53]

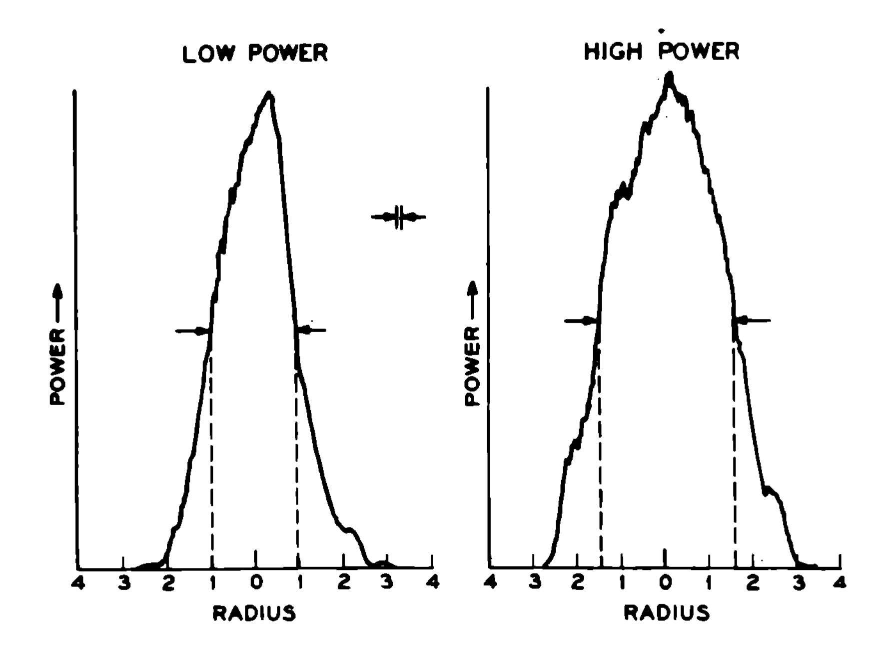


Figure 26 :
Scan of a near-field image at high and low powers 6 mm from the sample where the half-near-field distance is 2.3 cm. (Courtesy A. Ashkin). [53]

IV. PERSPECTIVE FOR THE FUTURE

Future developments will depend on the requirements of those applications which are most in demand. For the nearest future, the area of communications will experience the most active progress. The low cost of optical fibers and their large frequency bandwidth compared to those of wire lines makes optical-fiber communications an attractive substitute for present wired telephones. For this application, integrated optics will probably become the major switching means for selecting and interconnecting the various channels of communication. Here, light deflection will play an important part.

As to the spectral range that will be favored, one micron seems to be the immediately available region : optical fibers have lowest loss at about 1 μm ; the fastest directly modulated light sources are the GaAs LED and injection lasers operating at 0.9 μm ; high power Nd lasers emit at 1.06 μm ; single mode HeNe lasers have a line at 1.15 μm. Furthermore, silicon is a tailor-made detector for this spectral range. Si itself supports the most advanced integrated circuit technology and forms an ideal companion to integrated optics.

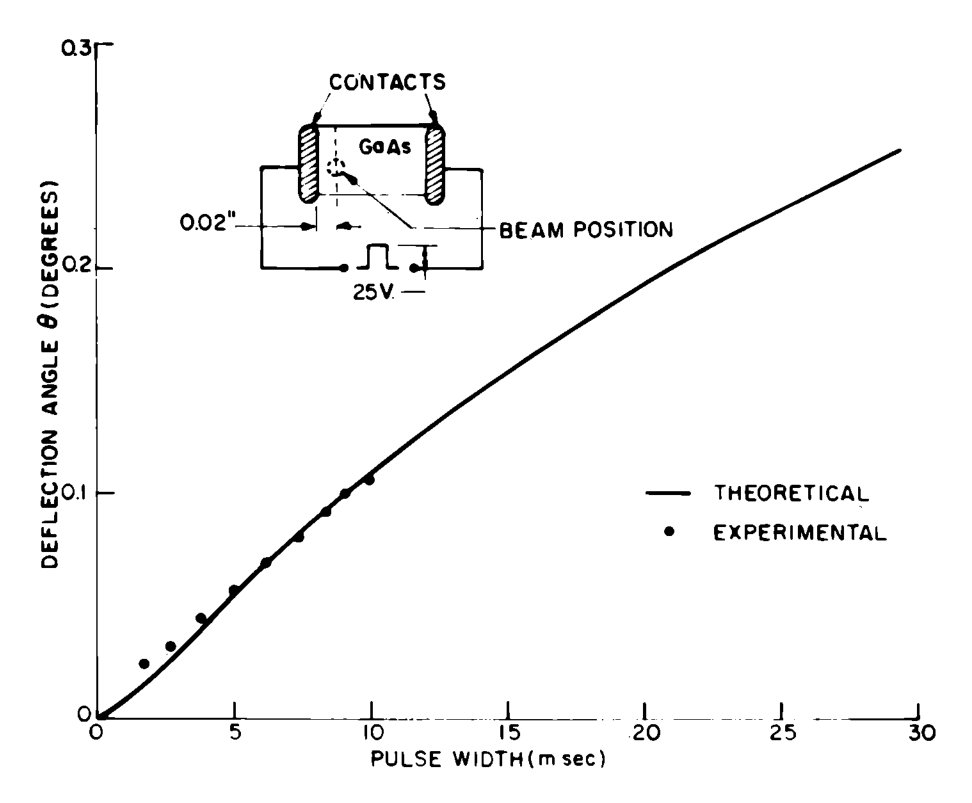


Figure 27 :
Pulse width dependence of beam deflection through a 0.0I inch GaAs sample. (Courtesy S.G. Liu). [11]

Major advances will require improvement in material homogeneity to reduce scattering and local changes in the rate of phase shift. In image display systems such defects degrade the quality of the image and the contrast ratio. Single crystals will be preferred to heteroepitaxial systems because of lower internal stresses. In heteroepitaxial layers, a slight structural mismatch may result in homogeneities . Changing the refractive index by diffusion will respect the crystalline properties by not introducing stresses. Lithium diffusion in LNT is a good case in point.

The future of magneto-optics is more cloudy. Its present marginal usefulness in the infrared may change with a breakthrough in materials. The control of domain spacing is certainly attractive. Buble technology may open new forms of optical data processing.

The optical equivalent of the RF directional coupler has been demonstrated by Somekh, et al. using integrated optics. [55] The structure consists of closely spaced parallel channels of higher

refractive index separated by lower index regions (Fig. 28). The spacing is small enough that the fringing field from one channel couples energy to adjacent channels. This coupling is cumulative if the waves travel in adjacent channels with the same phase velocity. In this experiment, after a travel of 2 mm all the light was transferred from the original channel into the adjacent channels. In the experimental structure, the distribution of the refractive-index was predetermined during the fabrication of the integrated-optics device : the high refractive index stripes were generated by proton bombardement through a mask in which high resolution openings were machined by ion beam etching. The coupling between channels would be controlled by applying an electric field to portions of each channel. Phase retardation by the transverse electro-optic effect was demonstrated in a separate experiment using a single channel. [56]

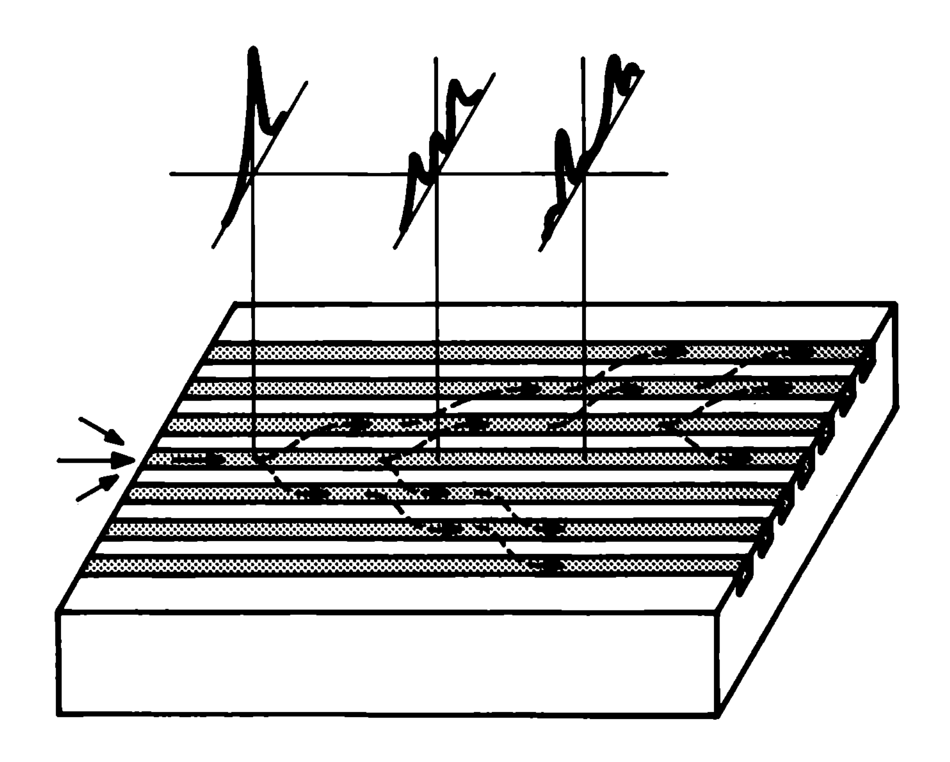


Figure 28 :
Optical waveguide directional coupler. The arrows show the flow of light into adjacent channels. The energy distribution at various positions down the coupler is sketched above the device. (After S. Somekh, et al.). [55]

Electro-optic deflection is now very practical for digital channel selection. It is a base-band system that can be driven with

a dc bias. In contrast, acousto-optic deflection requires the modulation of a carrier wave which drives the transducer. This inherently necessitates an additional stage of circuitry (admittedly not a large factor in the age of integrated circuits) but still an extra step in information processing. However, for continuous analog deflection, acousto-optics has unequaled advantages.

The potential uses of the acousto-optic effect to gain insight into the microstructure of specimens, expecially biological ones, is staggering. Consider the advent of acoustic microscopy which permits to probe specimens with soundwaves at a resolution comparable to that of optical microscopy : A diffraction limit of 1 μm has already been achieved. [57] By way of illustration, the acoustic wave sees a change in refractive index by a factor of 8 when it goes from sapphire to water. This large change in refraction permits the construction of a simple spherical objective lens. The specimen can be moved by a mechanical scanner at the focal point of the objective lens to map the acoustic image. The advantage of acoustic over optical or electron microscopy is that rather than examining a surface topograph, one examines changes in viscosity and hardness even below the surface of the specimen. With superfluid He, even greater resolution may be possible since sound velocity in this medium is very low (250 cm/sec) and wavelengths of 2 Å can be generated. [58]

▪ V. CONCLUSION

There are many principles to chose from for controlling the refraction of light inside a transparent material. The linear electro-optic or Pöckels effect has the advantage of high speed and efficiency. The acousto-optic effect may be more useful in restricted applications where continuous deflection of a light beam is needed.

Of all the materials available for integrated optics, GaAs and its alloys with AlAs will probably be the object of the most intensive exploration. The III-V compound semiconductor can provide all the basic optical functions needed for information processing : light generation, waveguiding, switching, coupling, modulation and even detection. However, future breakthroughs in material technology may drastically change our approach to light deflection.

▪ ACKNOWLEDGEMENTS

I am grateful to J. Alexander and M. Harvey for fabricating the GaAs p-n junction deflectors, to J.E. Berkeyheiser for technical assistance, to Drs. A. Ashkin, J.H. Mcfee, and G.E. Fenner for kindly supplying original figures from their work and to L Gorog, J. Hammer, J.D. Knox, and A. Miller for tutorial sessions.

■ REFERENCES

[1] F. Stern, Solid State Physics, F. Seitz and D. Turnbull, eds., Academic Press Inc., Vol 15, p. 327, (1963)

[2] J.H. Carran, L.A. D'Asaro, J.C. Dyment, and G.J. Herskowitz, I.E.E.E. J. Quantum Electronics QE-6, 367 (1970).

[3] H.Y. Fan, W.G. Spitzer, and R.J. Collins, Phys. Rev. 101, 566 (1956)

[4] E. Burstein, Phys. Rev. 93 , 632 (1954).

[5] T.S. Moss, Proc. Phys. Soc. (London) B 76 , 775 (1954).

[6] J.I. Pankove, Optical Processes in Semiconductors, Prentice Hall (1971) , pp. 10-12, 43.

[7] D. Redfield and M.A. Afromowitz, Appl. Phys. Letters 11 ,138 (1967).

[8] D.D. Sell, H.C. Casey, Jr., and K.W. Wecht, J. Appl. Phys. 45 , 2650 (1974). I am grateful to H.C. Casey for making this data available in advance of publication.

[9] E.M. Pell, Phys. Rev. 119 , 1222 (1960).

[10] O.V. Bogdankevich, Private communication.

[11] S.G. Liu and W.L. Walters, Proc. I.E.E.E. 53 , 522 (1965).

[12] F.R. Nash, J. Appl. Phys. 44 , 4696 (1973).

[13] J. Feinleib, S. Groves, W. Paul, and R. Zallen, Phys. Rev. 131 , 2070 (1963)

[14] F. Stern, Phys. Rev. 133 , A1,653 (1964).

[15] W. Franz, Z. Naturf. 13a , 484 (1958).

[16] L.V. Keldysh, Soviet Physics JETP 7, 788 (1958).

[17] V.S. Vavilov and K.I. Britsyn, Sov. Phys. Solid. State 2 1746 (1961).

[18] T.S. Moss, J. Appl. Phys. Supplement 32 , 2136 (1961)

[19] A. Frova and P. Handler, Phys. Rev. 137 , A 1857 (1965).

[20] B.O. Serafin and N. Bottka, Bull. Am. Phys. Soc. 10 106 (1965)

[21] K.W. Böer, Phys. Stat. Sol 9 , K35 (1965).

[22] I.P. Kaminow, I.E.E.E. J. Quantum Electronics 4, 23 (1968).

[23] S. Namba, J. Opt. Soc. Am. 51 , 76 (1961).

[24] J.E. Kiefer and A. Yariv, Appl. Phys. Letters 15 , 26 (1969).

[25] R.Y. Chiao, E. Garmire, and C.H. Townes, Phys. Rev. Letters 13 , 479 (1964).

[26] P.L. Kelley, Phys. Rev. Letters 15 , 1005 (1965).

[27] A. Javan and P.L. Kelley, I.E.E.E. J. Quantum Electronics QE-2 , 470 (1966)

[28] T.C. Lee and J.D. Zook, I.E.E.E. J. Quantum Electronics QE-4 , 442 (1968).

[29] J.M. Hammer and W. Phillips, Appl. Phys. Letters 24, 545 (1974).

[30] J.M. Hammer, D.J. Channin, and M.T. Duffy, Appl. Phys. Letters 23 , 176 (1973).

[31] J.M. Hammer, private communication.

[32] L. Brillouin , Ann. Phys. (Paris) 17 , 88 (1922).

[33] R. Lucas and P. Biquard, J. Phys. Rad. 7th Ser. 3, 464 (1932).

[34] P. Debye and F.W. Sears, Proc. Nat'l . Acad. Sci. (USA) 18 409 (1932).

[35] D.A. Pinnow , I.E.E.E. J. Quantum Electronics QE-6, 223 (1970).

[36] N. Uchida and Y. Ohmachi, Japan. J. Appl. Phys. 9, 155 (1970).

[37] E.I. Gordon, Appl. Optics 5 , 1629 (1966).

[38] A. Korpel, R. Adler, P. Desmares, and W. Watson, Appl. Optics 5, 1667 (1966).

[39] I. Gorog, J.D. Knox, P.V. Goedertier, and I. Shidlovsky, RCA Review 33 , 667 (1972) and R.W. Cohen, I. Gorog, and J.D. Knox, paper presented at the 1974 Spring Meeting of the Opt. Soc. of Am., Washington, D.C., April 21-25, 1974.

[40] R. Adler, I.E.E.E. Spectrum 4, 42 (1967).

[41] P.O. Sliva and R. Bray, Phys. Rev. Letters 14 , 372 (1965)

[42] C.S. Kumar, P.O. Sliva, and R. Bray, Phys. Rev. 169 , 680 (1968).

[43] A.R. Moore, Appl . Phys. Letters 13 , 126 (1968).

[44] B.W. Hakki, Appl. Phys. Letters 11 , 153 (1967).

[45] J.I. Pankove and A.R. Moore , R.C.A. Review 30 , 53 (1969).

[46] Y. Nannichi and I. Sakuma, J. Appl. Phys. 40 , 3063 (1969).

[47] J.P. Powers, Acoustical Holography, Vol 4, p. 533, G. Wade ed., Plenum (1972).

[48] R.K. Mueller and P.N. Keating, Acoustical Holography, Vol 1, p. 49, A.F. Metherel , et al., eds , Plenum (1969).

[49] T.R. Johansen, D.I. Norman, and E.J. Torok, J. Appl. Phys. 42, 1715 (1971).

[50] G.E. Fenner, U.S. Patent 3, 436, 679 and unpublished paper presented at 1966 IEEE. Solid State Device Research Conference.

[51] H. Rupprecht, J.M. Woodall, K. Konnerth, and D.G. Pettit, Appl. Phys. Letters 9, 221 (1966).

[52] J.H. McFee, R.E. Nahory, M.A. Pollack, and R.A. Logan, Appl. Phys. Letters 23 , 571 (1973).

[53] A. Ashkin, B. Tell, and J.M. Dziedzic , I.E.E.E. J. Quantum Electronics QE-3, 400 (1967).

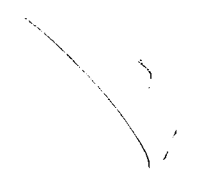
[54] F.W. Dabby and J.R. Whinnery, Appl. Phys. Letters 13 , 284 (1968).

[55] S. Somekh, E. Garmire, A. Yariv, H.L. Garvin, and R.G. Hunsperger Appl. Phys. Letters 22 , 46 (1973).

[56] D. Hall, A. Yariv, and E. Garmire, Optics Communications 1, 403 (1970).

[57] C.F. Quate, Private communication.

[58] C.H. Anderson, Private communication.



MAGNETOOPTICAL EFFECTS AND MATERIALS

Henri Le Gall

Laboratoire de Magnétisme et d'Optique

des Solides CNRS 92190 Meudon-Bellevue

France

I. INTRODUCTION

The experimental and theoretical studies of the spin-photon interactions are in constant increase since the first observation of the Faraday effect in 1845. Since this date, many magnetooptical effects have been discovered like the magnetic Kerr effect (1876), the normal and anomalous Zeeman effects (1896) and the Cotton-Mouton or Voigt effect (1906). These usual effects observed since about one century are fundamental for the microscopic study of the ions levels in gas, liquid and solid phases. By using the laser technique, new spin-photon interactions have been observed in the recent past (1961). These new processes may be separated into three groups.

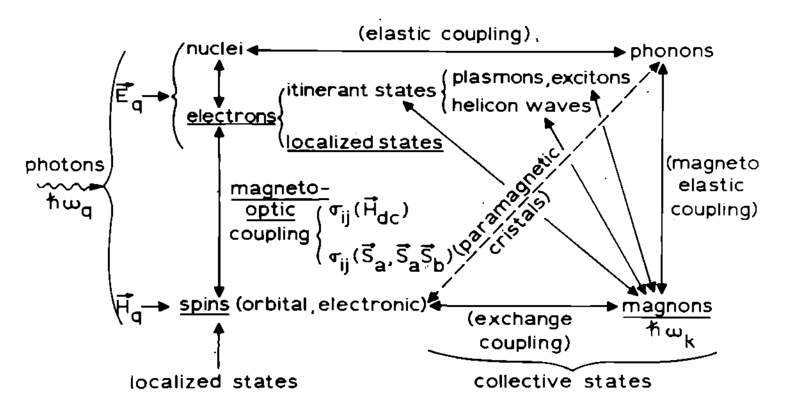
In the first group we have the 2-magnon absorption first observed in the far infrared in the antiferromagnetic fluoride FeF_2 and the 2-magnon- 1-phonon absorption discovered in a ferromagnetic crystal of nickel oxyde and in the antiferromagnet K Ni F_3. In these processes one photon is destroyed and 2 magnons, or 2 magnons plus one phonon, are created.

The second group corresponds to the magnon-sidebands or the exciton-magnon interactions. In these processes one photon may be absorbed or emitted and one magnon plus one exciton are created or destroyed.

The third group corresponds to the spin-Raman scattering. In ferro-ferri- and antiferromagnets the spin-Raman effect becomes the 1, 2 and 4-magnon Raman effects where the scattering of one photon produces the creation or the destruction of one, two or four magnons in the Stokes and the anti-Stokes processes respectively.

As shown in the Figure 1 the spin-photon interactions are induced either by a direct process produced by the magnetic dipole coupling or by an indirect process produced by the electric dipole coupling. The energy transfer from the electromagnetic field to the different localized and collective states of the magnetic crystal arises through the magnetic, electric, exchange, magnetooptical, photoelastic

magnetoelastic couplings. The spin-photon or magnetooptical (M.O.) interactions discussed above are obtained from these various couplings.



— Energy transfer from the E.M. field to the different states of a magnetic crystal.

Figure 1

Classifications of the spin-photon interactions may be obtained from general properties. The magnetooptical interactions may be separated either between the magnetic and the electric dipole transitions, or between scattering and absorption processes or between interactions via one or two magnetic ions. All these three groups may be divided again into two classes associated with the first and the second order magnetooptical effects.

This appears clearly in the table 1 where the second order processes, induced by the electric and the magnetic dipole couplings, are separated between the one and two magnetic ions models. Since the magnetooptical couplings are small the permeability and permittivity tensors can be expanded in increasing powers of the magnetization. The coefficients in the expansions are directly associated with the first-order, second order magnetooptical constants and so on.....

In the microwave range, the perpendicular and parallel pumpings are induced respectively by the first and the second-order magnetic transitions. On the other hand the para- and ferromagnetic resonances correspond to the one magnetic ion interactions whereas the ferri, antiferro and exchange resonances imply spin-photon interaction via 2 magnetic ions.

In the optical range we note that the Faraday effect is a first order interaction induced via the spin-orbit coupling of the excited orbital state of one ion whereas the Cotton-Mouton effect may be induced either via the spin-orbit coupling of the excited

SPIN-PHOTON INTERACTIONS		1st ORDER	2d ORDER
MAGNETIC DIP.TR.	$\mu_{ij} = \mu_{ij}^{(o)}$	$+ \; a_{ijk} M_k$	$+ \; a_{ijk\ell} M_k M_\ell + \dots$
ELECTRIC DIP.TR.	$\varepsilon_{ij} = \varepsilon_{ij}^{(o)}$	$+ \; b_{ijk} M_k$	$+ \; b_{ijk\ell} M_k M_\ell + \dots$
		⊥ R.F. PUMPING (PARA. AND FERRO. RESONANCE)	// R.F. PUMPING
SPIN-PHOTON INTERACTION WITH	1 - ION (VIA $\lambda \vec{L}_a \vec{S}_a$ INTERACTION)	FARADAY ($L_a^z S_a^z$) R1 ($L_a^\pm S_a^\pm$)	C.M. $(L_a^z S_a^z)^2$ R2 $(L_a^\pm S_a^\pm)^2$ R1 $(L_a^z S_a^z L_a^+ S_a^-)$
	2 - IONS (VIA EXCHANGE INTERACTION $J_{ab}^{ij} S_a^i S_b^j$)	R.F. AND OPTICAL PUMPING (FERRI, A.F. AND EXCH. RESONANCE) PHOTOMAGNETISM	CM $(S_a^z S_b^z)$ R1 $(S_a^z S_b^x)$ R2 $(S_a^x S_b^x)$ 2M - ABS. 2M + 1PH - ABS. EXCITON-MAGNON ABS. AND FLUOR.

CM = COTTON MOUTON EFFECT
R1,2 = 1-2 MAGNON RAMAN EFFECTS
M = MAGNON
PH = PHONON

Table 1

orbital sate of one ion, or via the exchange coupling of the excited state of a pair of magnetic ions. Effects observed recently like the 2-magnon absorption, the exciton-magnon absorption and the photomagnetism, which produces charge transfers, are explained only from a two-ion model.

One interesting classification of the spin-photon interaction is obtained by separating the absorption and the emission processes from the dispersive processes which produce the scattering of the light by the magnetic crystal as shown in the table 2. The physical and analytical connections between most of the spin-photon interactions may be discussed from two points of view. First, the well-known Kramers-Kronig's relations connect scattering and absorption processes like the Faraday and the Cotton-Mouton effects with the circular and linear magnetic dichroïsms. Recently we gave a second class of relations between the dispersive phenomena. These interactions may be separated between the elastic and the inelastic spin-photon scattering which relate, at the first-order the Faraday and the one-magnon Raman effects, and at the second-order the Cot-Cotton-Mouton and the one and the two-magnon Raman effects. It is to be noted that the one-magnon Raman effect may be induced by the second-order magnetooptical interaction via the $M_Z M_X$ term which is a linear function of the magnon operator.

Finally we observe, that the dispersive part of the magnetic resonance which appears in the microwave range, may produce elastic

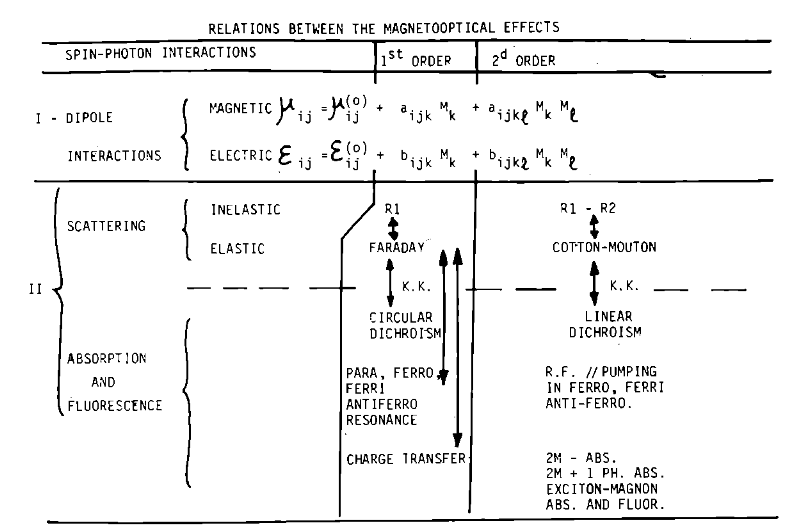


Table 2

and inelastic spin-photon scattering in the optical range as observed in ferrimagnetic garnets from the Faraday and the one-magnon Raman experiments. As shown in the Figure 2, the crystal has bigytropic properties in this range. In this figure the interactions of the electromagnetic field with the magnetic crystal are presented as a function of the wavelength.

The activity in magnetooptics may be divided into two groups. In the first one, the magnetooptical properties of crystals are analyzed in order to give informations, for example on the origin of the spin-photon interaction. In the second group, the magnetooptical effects are used as a tool to investigate optical and magnetic properties of crystals. The circular and linear magnetic birefringences and dichroisms are often used for the study of the atomic and molecular structure in gas, liquids and solids, whereas the two-magnon absorption, the exciton-magnon interaction and the magnon-Raman scattering give many details about the edge of the magnon Brillouin-zone and the density of state of the magnon.

▪ II. MACROSCOPIC DESCRIPTION OF THE MAGNETOOPTICAL EFFECTS : THE DIPOLE INDUCED BY THE LIGHT.

The macroscopic classical or quantum description of the MO effects is obtained without difficulty by scattering from the electric $\vec{p}_q$ and magnetic $\vec{m}_q$ dipoles induced in a magnetic crystal by the incident radiation of frequency ω_q. The electric dipole is deduced from the permittivity tensor which may be separated into one hermitian part ε^H_{ij} and one anti-hermitian part ε^A_{ij} associated respectively with the dispersive and

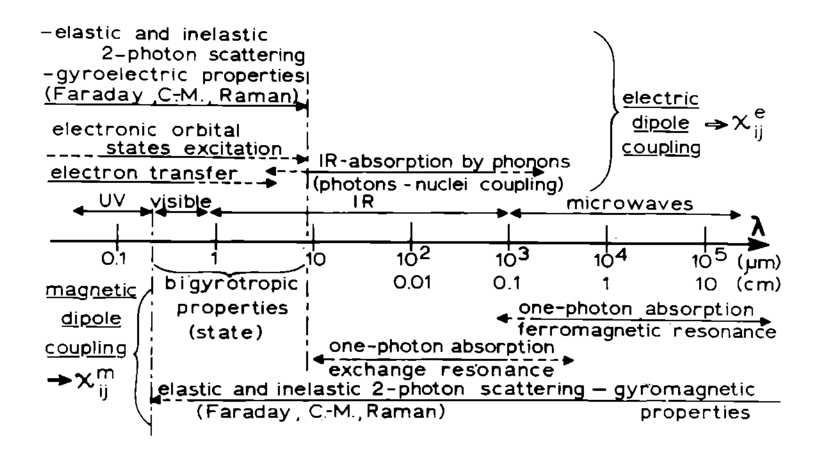


Figure 2

absorption phenomena. In what follows the study will be restricted to the dispersive processes (it is assumed a quasi transparent crystal : $\varepsilon^A_{ij} \simeq 0$). Therefore ε_{ij} is reduced to its hermitian part which is generally complex ($\varepsilon_{ij} = \varepsilon^H_{ij} = \varepsilon'_{ij} + i\varepsilon''_{ij}$). From the properties of symmetry of the spin-dependent permittivity tensor ($\varepsilon_{ij}(\vec{M}) = \varepsilon_{ji}(-\vec{M})$) it is shown without difficulty that, ε' and ε'' represent respectively the even and odd M.O. effects. The components in the complex plane of the electric dipole induced by the light in a magnetic crystal is deduced from the permittivity tensor such as :

$$\vec{P} = \sum_q \vec{P}^+_q \; e^{-i\omega_q t} + \vec{P}^-_q \; e^{+i\omega_q t}$$

$$\vec{P}^{\pm}_i = \mp \frac{i f^e_1}{4\pi} (\vec{E}^{\pm}_q \times \vec{M})_i + \frac{1}{4\pi} \sum_{j\ell m} f^{\ell m}_{2ij} M_\ell M_m E^{\pm}_j \qquad (1)$$

where i, j, ℓ,m =x, y, z and f^e_1 and $f^{\ell m}_{2ij}$ are the first- and second-order M.O. constants. The first-order M.O. interaction is described in the equation 1 by the cross product which represents the gyroelectric properties of the material. The second order M.O. interactions are described by a fourth rank tensor $f^{\ell m}_{ij}$ having 81 components. In the case of a

cubic crystal, the second-order effects are described by only three independent M.O. coefficient f_{11}, f_{12} and f_{44} (Voigt notation) such as :

$$\begin{vmatrix} \varepsilon'_{11} \\ \varepsilon'_{22} \\ \varepsilon'_{33} \\ \varepsilon'_{12} \\ \varepsilon'_{13} \\ \varepsilon'_{23} \\ \varepsilon'_{21} \\ \varepsilon'_{31} \\ \varepsilon'_{32} \end{vmatrix} = \begin{vmatrix} f_{11} & f_{12} & f_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ f_{12} & f_{11} & f_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ f_{12} & f_{12} & f_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & f_{44} & 0 & 0 & f_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{44} & 0 & 0 & f_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & f_{44} & 0 & 0 & f_{44} \\ 0 & 0 & 0 & f_{44} & 0 & 0 & f_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{44} & 0 & 0 & f_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & f_{44} & 0 & 0 & f_{44} \end{vmatrix} \begin{vmatrix} M_1 & M_1 \\ M_2 & M_2 \\ M_3 & M_3 \\ M_1 & M_2 \\ M_1 & M_3 \\ M_2 & M_3 \\ M_2 & M_1 \\ M_3 & M_1 \\ M_3 & M_2 \end{vmatrix}$$

A vectorial expression of the electric dipole is :

$$P_i^{\pm}\ (2) = \frac{1}{4\pi}\left[f_{12}\, \vec{M}^2\, E_i^{\pm} + 2f_{44} (\vec{M}.\vec{E}^{\pm}).M_i + \Delta f\, M_i^2\, E_i^{\pm} \right] \quad (2)$$

The first term is independent of the direction of the magnetization with respect to the electric field of the radiation. This is not the case for the second term which is maximum when $\vec{M}$ is parallel to $\vec{E}$. The third term represents the anisotropy of the second-order M.O. effects with $\Delta f = f_{11} - f_{12} - 2f_{44}$. The second order effects are isotropic when $\Delta f=0$. In that case the crystal is uniaxial with the optical axis parallel to $\vec{M}$. When $\Delta f \neq 0$ the second-order effects are anisotropic and the crystal becomes biaxial. An anisotropy coefficient ρ is defined :

$$\rho = 2f_{44} / (f_{11} - f_{12})$$

From the equation of evolution of the total magnetization $\vec{M} = \vec{M}_0 + \vec{m}_k + \vec{m}_q$ of the crystal submitted to the total magnetic field

$\vec{H} = \vec{H}_0 - \alpha\vec{M} + \vec{h}_q$, where M_0 and m_k are the static and the spin-wave magnetizations and α is the molecular field constant due to the exchange interaction, it is shown that the radiation $\vec{h}_q$ induces also in the optical band a magnetic dipole $\vec{m}_q$ (having a dispersive character) having in the complex plane the following components :

$$\vec{m}^{\pm}_{q,k} = \pm \frac{i f_1^m}{4\pi\varepsilon_r} (\vec{h}^{\pm}_q \times \vec{M}_k) \tag{3}$$

with $\vec{M}_k = \vec{M}_0 + \vec{m}_k$, $f_1^m = 4\pi|\gamma|\varepsilon_r\omega^{-1}_{q,h}$ and $\omega_{q,k} = \omega_q$ or $\omega_q \pm \omega_k$.

The magnetic susceptibility tensor can be deduced from this dipole. The cross product in the equation 3 shows the gyromagnetic properties of the material in the optical range.

A classical description of the M.O. interactions is obtained from the study of the electromagnetic field propagation in a magnetic crystal. That can be performed by introducing the electric and magnetic dipoles in the Maxwell equations. By assuming $\Delta f=0$, we can write :

$$\begin{aligned} \vec{q} \times \vec{E}^{-} &= \frac{\omega_q}{c} \left[\vec{h}^{-} - i\, \varepsilon_r^{-1}\, f_1^m\, (\vec{h}^{-} \times \vec{M}) \right] \\ \vec{q} \times \vec{h}^{-} &= -\frac{\omega_q}{c} \left[\varepsilon_r \vec{E}^{-} + i f_2^e\, (\vec{E}^{-} \times \vec{M}) + f_{12}\, \vec{M}^2\, \vec{E}^{-} + 2 f_{44}\, (\vec{M}.\vec{E}).\, \vec{M} \right] \end{aligned} \tag{4}$$

The eigen modes of propagation can be derived without difficulty for only two cases corresponding to the Faraday geometry (the wave vector $\vec{q}$ parallel to the magnetization $\vec{M}$) and the Cotton-Mouton geometry ($\vec{q} \perp \vec{M}$) such as :

$$\vec{q}_{/\!/}\vec{M} \rightarrow q_{\substack{\alpha\\ \beta}} \sim q_0 \left[\varepsilon_r \pm (f_1^m - f_1^e)\, M_z \right]^{1/2} = q_{\substack{r\\ \ell}}$$

$$\vec{q}_{\perp}\vec{M} \rightarrow \begin{cases} q_\alpha \sim q_0 \left| \varepsilon_r + f_{12}\, .M_z^2 \right|^{1/2} = q_{\perp} \\ q_\beta \sim q_0 \left[\varepsilon_r + (f_{12} + 2f_{44})\, M_z^2 \right]^{1/2} = q_{/\!/} \end{cases}$$

There are two eigen modes in each case. In the first case the eigen modes are left and right circular polarized, whereas in the second case they are linearly polarized parallel and perpendicular to the magnetization. That corresponds to the circular and linear magnetic birefringences described by the following Faraday and Cotton-Mouton specific phase-shifts :

$$\begin{aligned} \phi_F &= \frac{1}{2}(q_\ell - q_r) = \phi^e + \phi^m = -\frac{q}{2nc}(f_1^e - f_1^m)\, M_z \\ \phi_{CM} &= (q_{/\!/} - q_{\perp}) = \frac{\omega_q}{nc} f_{44}\, M_z^2 \end{aligned} \tag{5}$$

■ III. QUANTUM MACROSCOPICAL DESCRIPTION

A quantum macroscopic description of the M.O. effects can be obtained from the electromagnetic energy density in the crystal :

$$W = \frac{1}{8\pi}(\vec{E}\cdot\vec{D} + \vec{h}\cdot\vec{B})$$

such as the electric and magnetic dipole interaction hamiltonian densities are given by :

$$H^e = -\frac{1}{2}\sum_q (\vec{E}_q^+ \vec{P}_q^- + \vec{E}_q^- \vec{P}_q^+)$$

$$M^m = -\frac{1}{2}\sum_q (\vec{h}_q^+ \vec{m}_q^- + \vec{h}_q^- \vec{m}_q^+)$$

where E^- and P^- are in the complex plane the half of the Fresnel vector of the electric field and the electric dipole . E^+ and P^+ are the complex conjugate of E^- and P^-. It is interesting to express the hamiltonian with these quantities when we have to make a transfer from the classical to the quantum description. Indeed with the quantum description E^+ and E^- (or h^+ and h^-) are the destruction b_q and the creation b_q^+ photon operators, whereas $m^\pm$ which are functions of $M^\pm = M_x \pm iM_y$ can be expressed with spin-wave or magnon operators a_k and a_k^+ . By introducing the photon and magnon operators in H^e and H^m we obtain the hamiltonians of the elastic and inelastic spin-photon scattering induced by the electric and magnetic dipole interactions. Without describing all the M.O. processes we can discuss for example the hamiltonian of the first order magnetooptical interaction induced by the electric dipole coupling:

$$\mathcal{H}^e(1) = \mathcal{H}^e_z(1) + \mathcal{H}^e_{xy}(1) \qquad (6)$$

with $\mathcal{H}^e_{xy}(1) = \mathcal{H}^e_S(1) + \mathcal{H}^e_{AS}(1)$

$$\mathcal{H}^e_z(1) = \sum_{1,2} -i\,F^z_{1z}\,(\vec{e}_{1x}\,\vec{e}_{2y} - \vec{e}_{1y}\,\vec{e}_{2x})\,b_1 b_2^+\,\Delta\,(\vec{q}_1 - \vec{q}_2) + \text{h.c.}$$

$$\mathcal{H}^e_S(1) = \sum_{q_1 q_2 k} F^x_{1x}\,|(\vec{e}_{1\pi}\,\vec{e}^{\,\star}_{2r} - \vec{e}_{1r}\,\vec{e}_{2\pi})\,b_1\,b_2^+\,a_K^+\,\Delta\,(\vec{q}_1 - \vec{q}_2 - \vec{k}) +$$

$$+(\vec{e}^{\,\star}_{1r}\,\vec{e}_{2\pi} - \vec{e}_{1\pi}\,\vec{e}_{2r})\,b_1^+\,b_2 a_k^+\,\Delta(\vec{q}_1 - \vec{q}_2 - \vec{k})$$

$$\mathcal{H}^e_{AS}(1) = \mathcal{H}^e_S(1)\ (r \to \ell \text{ and } a_k^+ \to a_k)$$

where F^z_{1z} and F^x_{1x} are coefficients depending on the refractive index, on f^e_1 and M_0.

The operator $b_1 b_2^+$ shows that the hamiltonian $\mathcal{H}^e_z(1)$ corresponds to an elastic spin-photon scattering which is the Faraday rotation defined by a new quantum phenomenological description. In such a process one incident photon polarized linearly along the ox direction is scattered in a perpendicular direction along oy. It is following a perpendicular direction along oy. It is shown that the calculation of the probabilities for these transitions gives the value of the specific Faraday rotation of the material. The operators $b_1 b_2^+ a_k^+$ and $b_1 b_2^+ a_k$ show that $\mathcal{H}_S$ and $\mathcal{H}_{AS}$ represent inelastic spin-photon scatterings corresponding respectively to the Stokes and the anti-Stokes one-magnon Raman effects. We define a one magnon Raman susceptibility χ_{R_1} from the scattered light power P_S and the probability W_R of these transitions in a sample of volume V :

$$P_S = \frac{1}{2}\,\omega_{q_1}\,\chi_{R_1}\,E^2_{q_1} = W_{R_1}\hbar\omega_{q_2}\,V^{-1}$$

In optically isotropic materials the Faraday F^z_{1z} and Raman F^x_{1x} coefficients are defined from the same linear M.O. constant f^e_1 . From that property and by using the Fermi golden rule it is straight forward to show that the one-magnon Raman susceptibility is a quadratic function of the specific Faraday rotation ϕ^e_F :

$$\chi^e_{R_1} = \left(\frac{c\lambda_0}{2\pi^2}\right)\phi^2_F \sum_k \frac{n_k g\mu_B}{\eta_k M_0 V}$$

when n_k and η_k are the population and the relaxation rate of the magnons $\hbar\omega_k$; λ_0 is the radiation wavelength in vacuum.

The general expression of the second-order M.O. interactions induced by the electric dipole interaction is obtained by introducing the dipole P(2) in the expression of H^e :

$$\mathcal{H}^e(2) = -\frac{1}{8\pi}\sum_{ij\ell m\, q_1 q_2} f^{\ell m}_{2ij}\, M_\ell\, M_m\, E^-_i(q_1)\, E^+_j(q_2) + \text{h.c.}$$

By introducing the photon and magnon operators it is easy to verify that three typical terms associated with M_z^2, $M_x M_y$ (or M_x^2 and M_y^2) and $M_z M_x$ (or $M_z M_y$) appear :

$M_z^2 f^{33}_{2zz,yy}$	$b_1 b_2^+$	elastic Cotton-Mouton scattering
	$b_1 b_2^+ a_1 a_2^+$	inelastic Cotton-Mouton scattering
$M_x M_y f^{xy}_{2ij}$	$b_1 b_2^+ a_1^+ a_2^+$	2 magnon Stokes scattering
	$b_1 b_2^+ a_1 a_2$	2 magnon anti-Stokes scattering

$$M_x^2 f_{2ij}^{xx} \begin{cases} b_1 b_2^+ a_1^+ a_2^+ & \text{idem} \quad \text{case } M_x M_y \\ b_1 b_2^+ a_1 a_2 & \\ b_1 b_2^+ a_1 a_2^+ & \text{idem} \quad \text{case 1} \end{cases}$$

$$M_z M_x f_{2ij}^{3x} \begin{cases} b_1 b_2^+ a_k^+ & \text{1-magnon Stokes scattering} \\ b_1 b_2^+ a_k & \text{1-magnon anti-Stokes scattering} \end{cases}$$

The full expressions of the hamiltonians associated with the cases given above can be derived without difficulty. In an optically isotropic medium, it is shown that the 2-magnon Raman susceptibility χ_{R_2} is a quadratic function of the specific Cotton-Mouton phase-shift ϕ_{CM}.

IV. MICROSCOPIC DESCRIPTION OF THE MAGNETOOPTICAL EFFECTS. -REAL AND VIRTUAL TRANSITIONS. -TENSORIAL DYNAMICAL POLARIZABILITY OF A MAGNETIC ION.

A simple microscopic description of the spin-photon interactions can be obtained from the real and virtual transitions which are at the origin of all the absorption and dispersive phenomena.

It is well known that the population change of an initial ground state $|i>$ (due to the absorption of one photon) or an initial excited state (due to radiative or non-radiative transition by emission of one photon or one phonon) produces the line-width Γ_i of this state (natural line-width) which is connected to the relaxation time T_i by the Heisenberg uncertainty relation : $\Gamma_i = \hbar/T_i$. The absorption or the emission of one photon is a real transition which arises only when the photon energy $\hbar\omega_q$ corresponds to the energy separation $\hbar\omega_{fi} = E_f - E_i$ between the initial and the final states such as :

$$|\hbar_q - \hbar\omega_{fi}| < \Gamma_i + \Gamma_f \qquad (7)$$

In the case of the relation 7 the photon is wholly absorbed. In the opposite case corresponding to the following relation :

$$|\hbar\omega_q - \hbar\omega_{fi}| > (\Gamma_i + \Gamma_f) \qquad (8)$$

due to the positive or negative gap $\Delta = \hbar(\omega_q - \omega_{fi})$, the photon cannot be absorbed. In that case the light-matter interaction still appear but through a *virtual transition* in which the photon is captured during a very short time Δt by the atom and then reemitted . This capture time

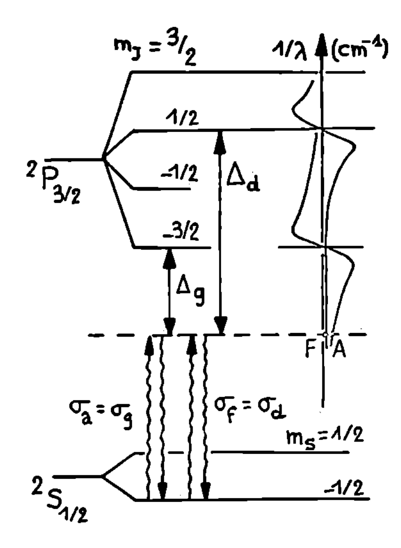
is also defined by the uncertainty relation of Heisenberg :

$$\Delta t = \frac{\hbar}{\Delta} \qquad (9)$$

The capture of the photon is fundamentally similar to the tunnel effect where Δ is the potential to be hoped by the whole system photon + atom. These transitions actually appear as two-step photon scatterings but they are referred to as "virtual" interactions because of their very short duration ($\Delta t \sim 10^{-16}$ second) compared to the usual relaxation times ($T = 10^{-5}$ to 10^{-8} sec) associated with the real transitions.

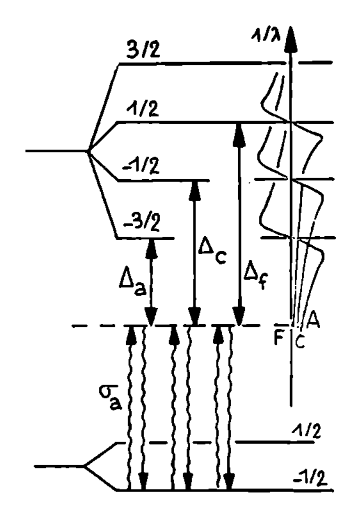
In a magnetic crystal (para-, ferro- or antiferromagnet) the dispersive effects like the circular and linear magnetic birefingences and the spin and magnon Raman scatterings are described by virtual transitions. On the other hand the absorption effects like the Zeeman effect, the circular and linear magnetic dichroisms, the 2-magnon absorption and fluorescence and the exciton-magnon transitions are due to real transitions.

The elastic spin-photon scatterings associated with the Faraday and Cotton-Mouton birefringences can be discussed from a Zeeman diagram corresponding for instance to the $^2S_{1/2} \rightarrow {}^2P_{3/2}$ transitions allowed by the usual selection rules as described in the figures 3 and 4. Only the lower level of the ground multiplet is assumed to be occupied. When the light is propagating along the direction of the magnetic field (oz axis), only the transitions σ_1 (or σ_{left}) and σ_2 (or σ_{right}) can be



Biréfringence Magnétique Circulaire

Figure 3



Birefringence Magnétique Linéaire

Figure 4

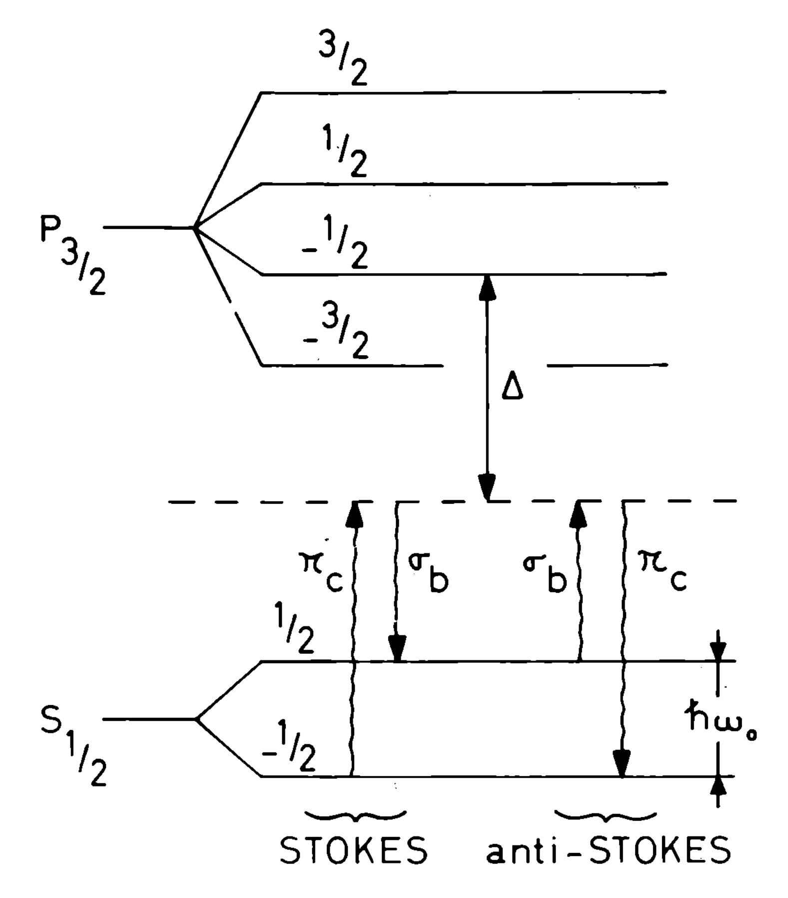


Figure 5

EFFET RAMAN DE SPIN

induced. In the case of virtual transitions the energy gaps Δ_ℓ and Δ_r associated with the left and right circular polarized photon are not the same, which produces a higher dispersion for the σ_ℓ component in the case of our figure and a rotation of the resultant linear polarization. The absorption difference between the two circular components gives the magnetic circular dichroism.

In the case of figure 4 the light is propagating perpendicularly to the d.c magnetic field direction and can induce either the σ_1 and σ_2 components when the incident light polarization E is perpendicular to the oz axis, or the σ_3 component when E is parallel to the oz axis. The difference of the energy gap in these two cases is at the origin of the two eigen modes of the electromagnetic propagation in the Cotton-Mouton configuration. When all the ground multiplet states are occupied ($T \neq 0°K$) we have to take into account, in the Faraday and Cotton-Mouton effects, all the transitions allowed by the selection rules.

The coupling of a ground state to an excited state by a virtual transition can, after the scattering, leave the magnetic ion in a different ground state. That is shown in figure 5 with the change of the magnetic quantum number M_S after the photon is reemitted. These virtual transitions can increase (Stokes component) or decrease (anti-Stokes component) the ground state energy by the quantity $\hbar\omega_0$. Such a magnetic change $\Delta M_S = \pm 1$ in the ground state is due to the spin-orbit coupling $\lambda' \vec{L}\vec{S}$ in the excited state which mixes the orbital and spin motions (or wave functions). Indeed the eigen functions of an excited state with spin-orbit coupling $|J\ M_J>$ may be expanded as a function of the unperturbated eigen functions $|M_L\ M_S>$ such as :

$$|L\ S\ J\ M_J> \;=\; \sum_{M_L M_S} |L\ S\ M_L\ M_S><L\ S\ M_L\ M_S\,|L\ S\ J\ M_J>$$

Therefore, during a virtual transition, the capture of one photon produces in a first time the connection of the ground state having the magnetic quantum number M_S to one of the excited states $|M_L\ \ M_S>$ contained in $|J\ M_J>$

In the second time the spin-orbit operator mixes or couples these excited states such as :

$$L^{\pm}\ \ S^{\mp}\ |M_L\ \ M_S> = \text{constant}\ \ |M_L \pm 1,\ \ M_S \mp 1>$$

Therefore we observe an inelastic scattering when, in a third time, one photon is reemitted with a transition between the excited state $|M_L \pm 1\ ,\ M_S \pm 1>$ and the ground state $|0,\ M_S \pm 1>$. That corresponds to the first-order spin Raman scattering. In the same way a virtual transition can produce the change $\Delta S = \pm 2$ in the ground state which corresponds to the second-order spin Raman scattering. In the ferro, ferri and antiferromagnets, due to the exchange coupling a localized excitation cannot arise. Therefore, if the virtual transition is localized on a single ion, the final magnetic excitation must be expanded on the N ionic spins of the crystal which have an averaged small variation $\pm 1/N$, $\pm 2/N$..... This corresponds to the 1 and 2-magnon Raman scattering. The preceding scatterings are induced by the electric dipole transitions. From the different levels of the ground multiplet, the magnetic dipole interactions can induce similar effects. The real transitions correspond to the usual spin resonances, and the virtual transitions between these ground levels can induce all the MO interactions discussed above.

1) Dynamical Polarizability of a magnetic ion.

The microscopic origins of the spin-photon interactions can be discussed quantitatively from the expression of the quantum dynamical polarizability of one magnetic ion or of a group of magnetic ions such as:

$$\chi^{+}_{ij} = \sum_{a} N_a\ \alpha^{+}_{ij}\ (a)$$

where $\alpha_{ij}(a)$ is one element of the dynamical polarizability tensor for the ion a. The total Hamiltonian which describes the magnetic crystal is given by :

$$H_1 = H_0 + H_R \quad \text{with}$$

$$H_0 = H_{c.f.} + H_{cr} + H_{ex} + H_{s0} + H_z$$

$$H_R = -\vec{E}\vec{P} = e\vec{E}\vec{r} = -e\,(\vec{E}^{+}\, e^{-i\omega_q t} + \vec{E}^{-}\, e^{i\omega_q t})\,(\vec{r}^{+}(t) + \vec{r}^{-}(t))$$

where the different terms of H_0 correspond to the evolution of the electronic system in the central field approximation (H_{cf}), the crystal field (H_{cr}), the spin-orbit interaction (H_{s0}) and the external magnetic field (H_z). The electromagnetic perturbation H_R induces a coupling between the states $|n>$ and the system now must be described by the perturbed states $|\Psi_n>$. The response of the magnetic system to an electromagnetic excitation is given by the dynamical polarizability $\alpha(a)$ of the magnetic ions which is deduced from the expression of the induced dipole $\vec{P}$ in the perturbed state $|\Psi_g>$ such as :

$$<\vec{P}> = \sum_g <\Psi_g|\vec{P}|\Psi_g>\rho_g = \sum_g <\Psi_g|e\vec{r}|\Psi_g>\rho_g \qquad (10)$$

The summation runs over all the levels of the ground multiplet having the occupation probability ρ_g. The components of the dipole are given by :

$$<P_i> = <P_i^+> + <P_i^-> \quad \text{with} \quad <P_i^{\pm}> = \sum_j \alpha_{ij}^{\pm}\, E_j^{\pm}\, e^{\mp i\omega_q t} \qquad (11)$$

The ground state perturbed by the radiation is obtained without difficulty from the time-dependent pertubation theory.

$$|\psi_g> = |g>\, e^{-i\omega_g t} + \frac{e}{\hbar}\sum_n \Big[<n|\vec{r}\vec{E}^{+}|g>\Big(\frac{e^{i(\omega_{ng}-\omega_q)t} - 1}{\omega_{ng} - \omega_q}\Big) +$$

$$+ <n|\vec{r}\vec{E}^{-}|g>\Big(\frac{e^{i(\omega_{ng}+\omega_q)t} - 1}{\omega_{ng} + \omega_q}\Big)\Big]|n> \qquad (12)$$

It is assumed that only the state $|g>$ is occupied at $t = 0$.
By introducing (12) in (10) it is shown that the dynamical polarization associated with E_J^+ is given by :

$$\alpha_{ij}^{+} = \frac{e^2}{\hbar}\sum_{g,n} \rho_g \{\mathrm{Re} <g|\vec{r}_i|n><n|\vec{r}_j|g>\Big[Z(\omega_{ng} - \omega_q) + Z(\omega_{ng} + \omega_q)\Big] +$$

$$+i\,\mathrm{Im} \langle g|\vec{r}_i\,|n\rangle\langle n|\vec{r}_j\,|g\rangle \left[Z(\omega_{ng} - \omega_q) - Z(\omega_{ng} + \omega_q)\right] \qquad (13)$$

with

$$Z^{*}(\omega_{ng}-\omega_q) \pm Z(\omega_{ng}+\omega_q) = \frac{P}{\omega_{ng}-\omega_q} \pm \frac{P}{\omega_{ng}+\omega_q} + i\pi\,\delta(\omega_{ng}-\omega_q) \mp i\pi\,\delta(\omega_{ng}+\omega_q) \qquad (14)$$

In fact the light-atom interaction is completely described by changing the electric dipole operator $\vec{r}_i$ by $\vec{R}_i$ such as :

$$\vec{R}_i = \left[\vec{r}_i + i(\vec{q}\cdot\vec{r})\cdot\vec{r} - \frac{\hbar}{2m\omega_q}(\vec{L} + 2\vec{S})\times\vec{q} + \dots\dots\right]_i \qquad (15)$$

where the second and third terms induce the electric quadrupole and magnetic dipole interactions respectively. These equations are general and valid for all the spin-photon interactions.

If the second state $|g\rangle$ is assumed to be only occupied at the initial time, the polarizability [13] associated with E^+ has 4 terms (see equation 14) which correspond to the real and virtual components of two transitions described in the Figure 6 by the matrix elements $\langle n|E^+ r^-|g\rangle$(transition 1) and $\langle g|E^+ r^+|n\rangle$ (transition 2). Transition 2 is always virtual, so that the fourth term $\delta(\omega_{ng} + \omega_q)$ vanishes always. The third term $\delta(\omega_{ng} - \omega_q)$ does not vanish when the transition 1 is resonant or real. The first and second terms describe the dispersion of the transitions 1 and 2 respectively. These terms produce the refraction, the natural and magnetic birefringences and the spin-Raman scattering.

Now we discuss the approximation case obtained far from the absorption bands where only the dispersive terms are non-zero. The relation 14 is restricted to :

$$\left.\begin{aligned} \frac{P}{\omega_{ng}-\omega_q} + \frac{P}{\omega_{ng}+\omega_q} &\simeq \frac{1}{\omega_{ng}-\omega_q} + \frac{1}{\omega_{ng}+\omega_q} = \frac{2\omega_{ng}}{\omega_{ng}^2 - \omega_q^2} \\ \frac{P}{\omega_{ng}-\omega_q} - \frac{P}{\omega_{ng}+\omega_q} &\simeq \frac{2}{\omega_{ng}^2 - \omega_q^2} \end{aligned}\right\} \qquad (16)$$

Due to the term ω_{ng} , these expressions introduce all the magnetic perturbations (spin- orbit, exchange and Zeeman interactions) which produce the elastic and inelastic spin-photon scattering. That can be shown by using the following simplifications (Figure 6-b):

$$|E_g^{(1)}|\,,|E_n(1)| \ll \hbar\omega_q \ll \hbar\omega_0 = E_n^0 - E_g^0$$



$E_g^{(1)}$ and $E_n^{(1)}$ correspond to the energy shifts of the unperturbed levels E_g^0 and E_n^0 produced by the crystal field, exchange, spin-orbit and Zeeman perturbations.

The frequency factor (14) or (16) of the Zeta Dirac function can be expanded such as :

$$\frac{P}{\omega_{ng} - \omega_q} - \frac{P}{\omega_{ng} + \omega_q} \simeq \frac{2}{\omega_{ng}^2 - \omega_q^2} \simeq \frac{2}{\omega_0^2 - \omega_q^2} \left[1 - \frac{2\omega_0 (E_0^{(1)} - E_g^{(1)})}{\hbar \, (\omega_0^2 - \omega_q^2)} + \ldots \right] \quad (17)$$

Therefore we obtain in the bracket increasing powers of $E_n^{(1)}$ and $E_g^{(1)}$ which are directly dependent on spin operators via the exchange and spin-orbit couplings such as :

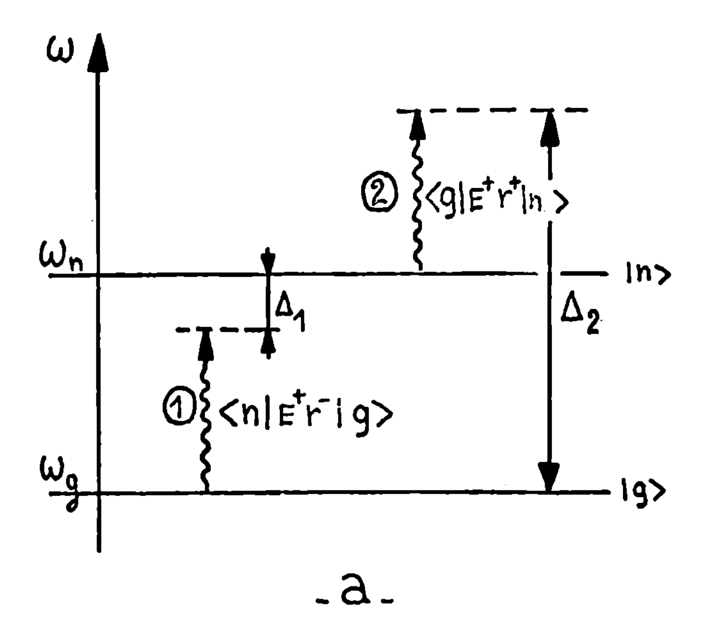
$$E_n^{(1)} = H_Z + H_{Cr} + \sum_a \lambda' \vec{L}_a \vec{S}_a + \sum_{a>b} J'_{ab} \vec{S}_a \vec{S}_b + \sum_{a>b} J''_{ab} \vec{S}_a \times \vec{S}_b \quad (18)$$

where λ' , J'_{ab} and J''_{ab} are the spin-orbit and exchange constants of the excited states. So we understand why the dielectric polarizability is dependent on the spins at the first-order, second-order and so on by introducing the equations (14), (17) and (18) in the equation(13). The calculation of the eigen functions $|g\rangle$ and $|n\rangle$ is not easy in a general case unless for some paramagnetic ions where it is possible to derive the expressions of the circular magnetic birefringences and dichroism. To conclude we give in the following table the matrix elements which describe the different MO effects with their perturbation order, MO order and the corresponding change of the magnetic state.

perturbations order	matrix elements of the transitions	effects	dif. type	MO order	$\frac{M_S}{g\mu_B}$
1	$\lvert\langle n\rvert r_j \lvert g\rangle\rvert^2$	Absorption 1 photon		0	0
2	$\langle g\rvert r_j^+ \lvert n\rangle\langle n\rvert r_j^- \lvert g\rangle$	Z.eff. $\sigma^+\sigma^-\pi$ el. scatt.			
	$\langle g'\rvert r_i^+ \lvert n'\rangle\langle n'\rvert H_{spin} \lvert n\rangle\langle n\rvert r_j^- \lvert g\rangle$		E*	0	0
3	$H_{spin} = \lambda' \vec{L}_a \vec{S}_a = \lambda' L_a^z S_a^z +$	Faraday	E	1	0
	$+ \frac{1}{2}(L_a^+ S_a^- + L_a^- S_a^+)$	1-magnon Raman	I**	1	±1
3	$H_{spin} = J' S_a S_b = J' S_a^z S_b^z +$ $+ \frac{1}{2}(S_a^+ S_b^- + S_a^- S_b^+)$	2-magnons Raman	I	2	0
	$(L_a^z S_a^z)^2$	Cotton-Mouton	E	2	0
4	$H_{spin} = (\lambda' \vec{L}_a \vec{S}_a)^2 \rightarrow (L_a^+ S_a^-)^2$	2-magnons Raman	I	2	±2
	$L_a^z S_a^z L_a^+ S_a^-$	1-magnon Raman	I	2	±1
4	$H_{spin} = (J' \vec{S}_a \vec{S}_b) \rightarrow (S_a^+ S_b^-)^2$	4-magnons Raman	I	4	0

* elastic

** inelastic



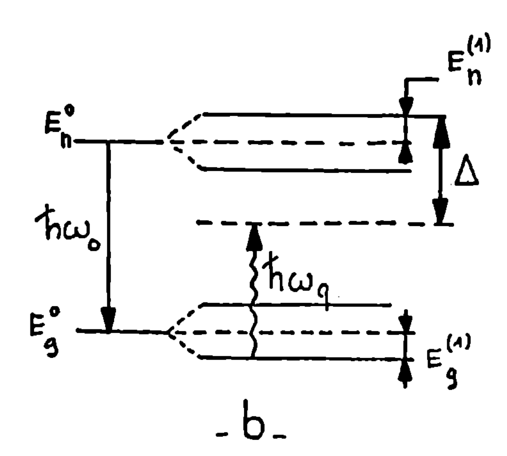


Figure 6

■ V MAGNETOOPTICAL MATERIALS

The first and second order M.O. interactions depend either on the amplitude and the direction of the magnetization with respect to the light propagation (Faraday effect), or on the amplitude of the magnetiza-

tion and the direction of light propagation with respect to the crystallographic axis (anisotropic Cotton-Mouton effect). All these properties are useful to control the light propagation (modulation, switch, gate, gyrators, isolator for example) with appropriate devices using bulk or thin-film (integrated optics) materials. On the figure 7 for instance, it is shown the basis of light modulation by Faraday effect. The magnetization reversal induced by the oscillating field of the modulation coil produces a corresponding time-dependent rotation of the light polarization. After a cross analyzer the light is amplitude modulated. The best materials for such devices must have high specific Faraday rotation ϕ_F and low absorption coefficient α, which correspond to the high merit factor ϕ_F/α. In integrated optical systems the light propagation depends also on the Cotton-Mouton effect.

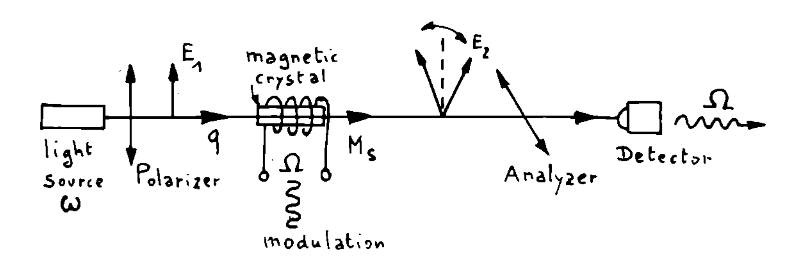


Figure 7

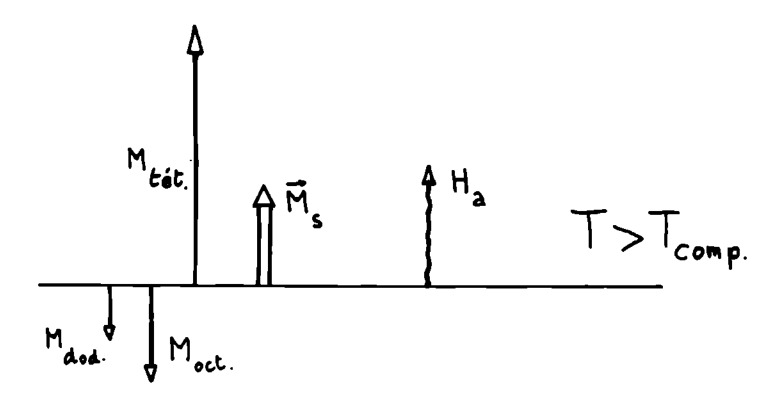
The Faraday and Cotton-Mouton effects can be important in ferro, ferri and weak ferromagnets. In antiferromagnet only the Cotton-Mouton effect can be detected. Among all the different types of magnetic crystal structures, the ferrimagnetic garnets are the best materials obtained until now for magnetooptical devices. Therefore in what follows the M.O. properties of the rare-earth ferrimagnetic garnets will be discussed in detail. The crystallographic structure of the garnets is cubic but magnetic and non-magnetic ions can occupy either the dodecahedral site (site C) (Y^{3+} , Bi^{3+} and all the rare earth ions from Pr^{3+} to Yb^{3+}) or the tetrahedral site (site d) (Fe^{3+}, Ga^{3+}, Al^{3+}.....) or the octahedral site (site a) (Fe^{3+}, Ru^{3+}, Ru^{4+}, Cr^{3+}, In^{3+}....). It can be shown that the Faraday rotations induced by the electric and magnetic transitions are proportional to the moment of the magnetic ions of the different sublattices such as for $T < T_{comp}$ where T_{comp} is the magnetic compensation temperature, we have :

$$\phi_F = \phi^e + \phi^m = -(A_e + A_m)\,|M_a| + (D_e + D_m)\,|M_d| - (C_e + C_m)\,|M_c| \qquad (19)$$

where A_e, D_e, C_e and A_m, D_m, C_m are the first-order M.O. coefficients associated with the electric and magnetic dipole transitions in the a,d and

c sites respectively. Al low temperatures the rare-earth moment is higher than the resulting moment of the Fe^{3+} ions in the a and d sites. Near the Néel temperature, an opposite situation is observed as shown in the figure 8, and all the signs of equation 19 must be inverted. It is to be noted that :

- all the magnetic coefficients A_m, D_m and C_m are positive.
- the electric and magnetic coefficients of a given site have opposite signs.



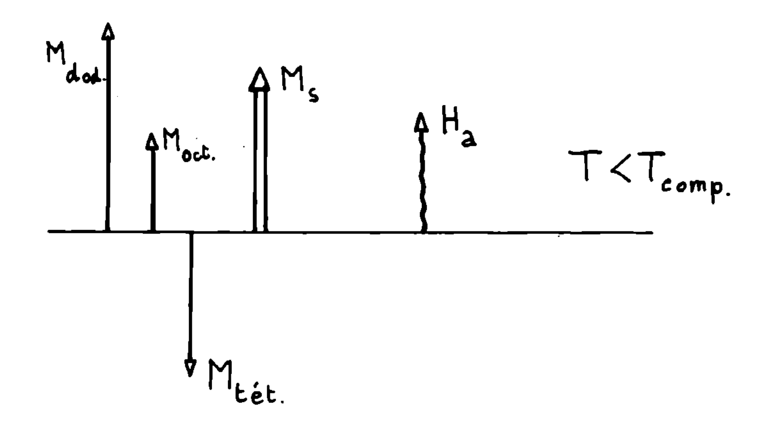


Figure 8

Starting from the simpler case of the ferrimagnetic yttrium-iron garnet (YIG), where the site c is occupied by the non-magnetic ions Y^{3+}, the Faraday rotation is described by :

$$\phi_F = \phi^e + \phi^m = -(A_e + A_m)\ |M_a| + (D_e + D_m)\ |M_d| \qquad (20)$$

M_a and M_d are measured in Bohr magneton μ_B per two formula units (in YIG there are 8 formula units $Y_3\ Fe_5\ O_{12}$ in a unit cell). The theoretical and experimental values of the magnetic coefficients are well-known ($A_m = D_m = 2\,\pi n\ c^{-1}\ \gamma = 9.15°/cm/\mu_B$ in YIG |1|) , so the "electric" Faraday rotation ϕ^e can be obtained from the total Faraday rotation as shown in the Figure 9 and equation 20. Therefore , by comparing the experimental value of

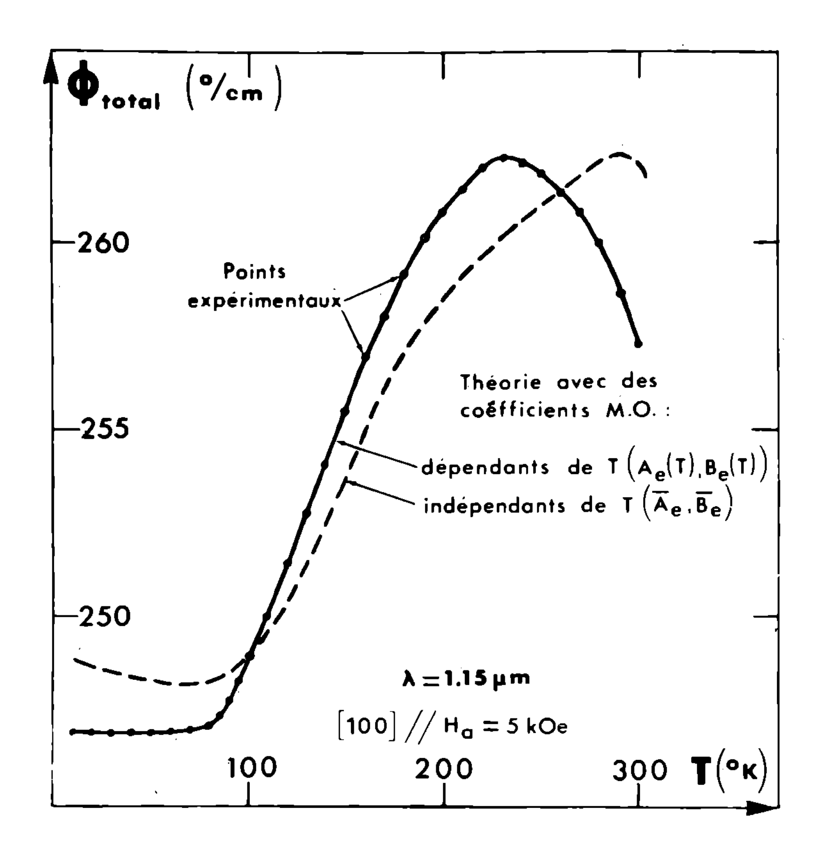


Figure 9

$\phi^e = -A_e\ |M_a| + D_e\ |M_d|$ with $|M_a|$ and $|M_d|$ for each temperature T we may deduce the temperature dependence of A_e and B_e as shown in the Figure 10 |2|. It is to be noted that the M.O. coefficients decrease from a maximum at low temperatures to zero near the Néel temperature and that for all temperatures in the ferrimagnetic phase D_e reaches approximately fifty per cent of A_e . $\bar{A}_e$ and $\bar{D}_e$ are the mean values obtained between 10 and 300°K by assuming the MO coefficients to be independent on temperature. These mean values depend on the temperature range of the experiment and therefore do not represent the intrinsic values of the M.O. coefficients. The description of the Faraday rotation in YIG can be summarized by the Figure 11

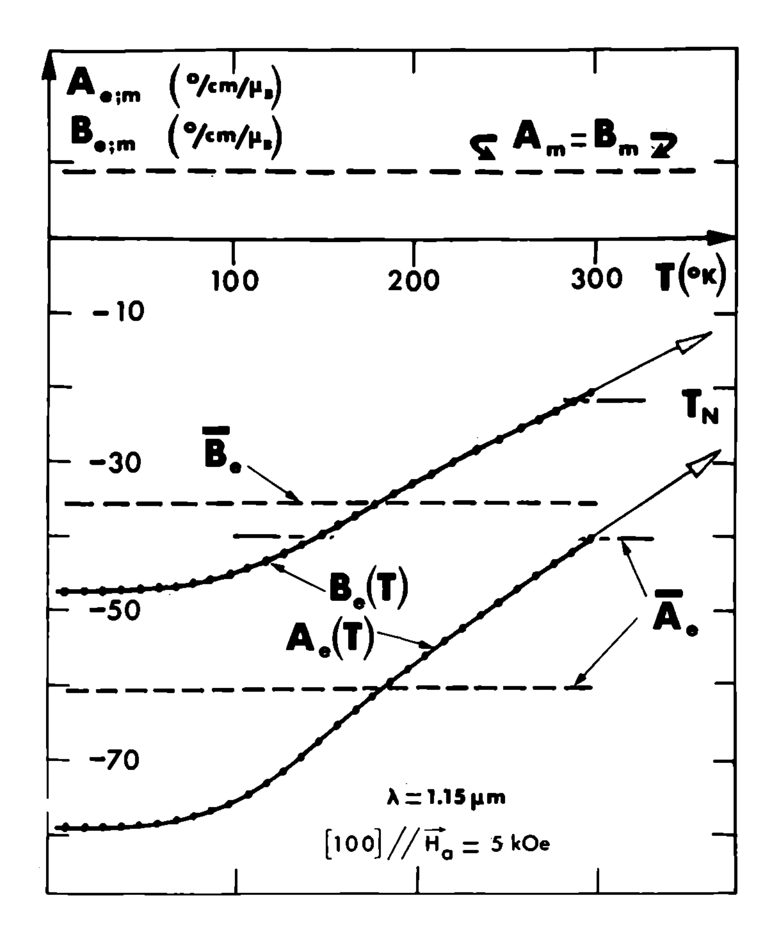


Figure 10

where the diameter of the circles are proportional to the electric (full circles) and magnetic (dotted circles) Faraday rotation in the a and d sites at room temperature. This figure suggests the possibility to increase the specific rotation by making appropriate substitution in the a, d and c sites. The magnetization is lower in the octahedral site but its contribution to the Faraday rotation is higher and opposite in sign to the rotation in the tetrahedral site. The substitution of Fe^{3+} in the d site by non magnetic ions as Ga^{3+} or Al^{3+} must increase the net rotation but the experiment does not confirm this hypothesis as shown in the Figure 12 where the Faraday rotation of $Y_3\ Fe_{5-x}\ Al_x\ O_{12}$ are given for different values of x (=0 to 1.34). It is suggested that the decrease of ϕ_F by increasing the substitution rate of Al ions is due to the corresponding decrease of the exchange interaction a-d. This situation must be compared with the decrease of the electric coefficients A_e and D_e (Figure 10) when T reaches the Néel temperature and which has been attributed to the parallel decrease of the molecular field.

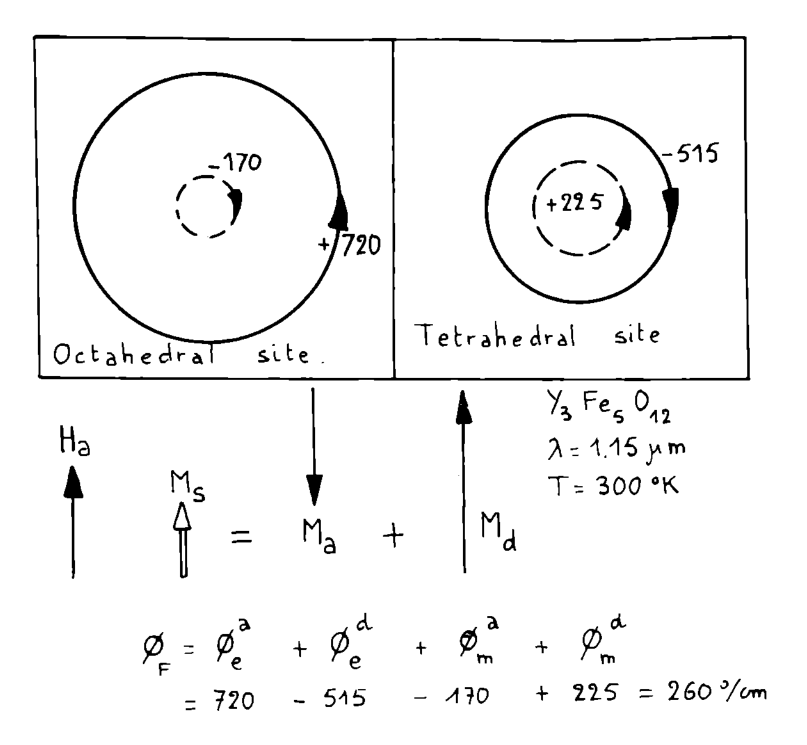


Figure 11

The second possibility to modify ϕ_F is obtained with the substitution of Y^{3+} by trivalent rare-earth ions in the dodecahedral site. Due to the strong spin-orbit couplings and high moments of the rare-earths, the corresponding contributions to ϕ_F are very important and parallel to the contribution of the sites d but antiparallel to the contribution of the sites a. The contribution of the rare-earth is very important mainly for $T < T_{comp}$ as shown in the Figure 13 for R_3 Fe_5 O_{12} (R= H ,D_y,Gd and Tb) garnets. Unfortunately so high specific rotation cannot be used at room temperature, but however it has been shown recently that YIG with small substitutions of Y^{3+} by Pr^{3+} and Nd^{3+} gives a good merit factor ϕ_F/α $|^3|$at the wavelength of the neodymium lasers (1.06 μm).

Finally the spin-orbit coupling of the rare-earth ions produces a strong magnetocrystalline anisotropy which requires high static magnetic fields to avoid domains structure and reach the saturation of the magnetization. A new type of ferrimagnetic garnets having anomalous large Faraday rotation and small magnetocrystalline anisotropy at room temperature has been studied recently by using the Bi-doped yttrium-iron garnets Y_{3-x} Bi_x Fe_5 O_{12}. By increasing the Bi content **x** the Faraday rotation decreases, its sign is changed and then increases strongly towards high negative values. The Faraday rotation spectra of pure YIG and Bi-doped YIG are compared in the

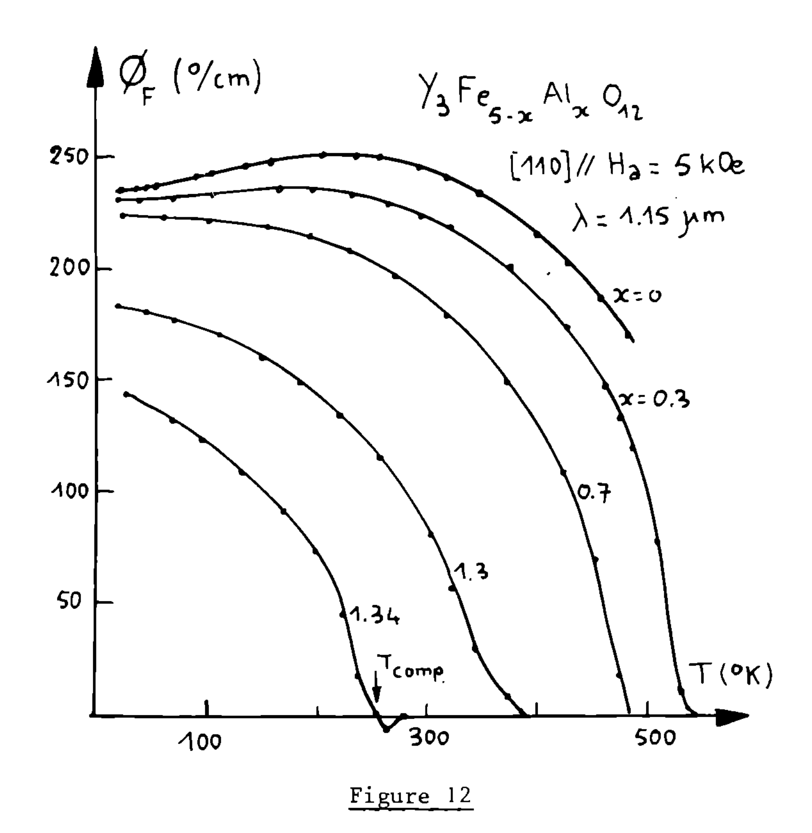


Figure 12

Figure 14 |4|. The large rotation has been attributed to a mixing of oxygen and bismuth orbitals which induces an enhancement of the spin-orbit coupling of the tetrahedral complex. Therefore the rotation of the tetrahedral complex which is negative is increasing with the Bi-content as indicated in the table 3.

λ (μm) → / x ↓	0.6328	1.15
0	+ 750	+ 260
0.07	-1450	- 150
0.22	-2000	- 340
0.3	-5900	- 420

Table 3

The propagation of light in magnetic materials can be strongly modified by the large values (ϕ_{CM}) of the Cotton-Mouton effect or linear magnetic birefringence (LMB). It has been observed that this second order

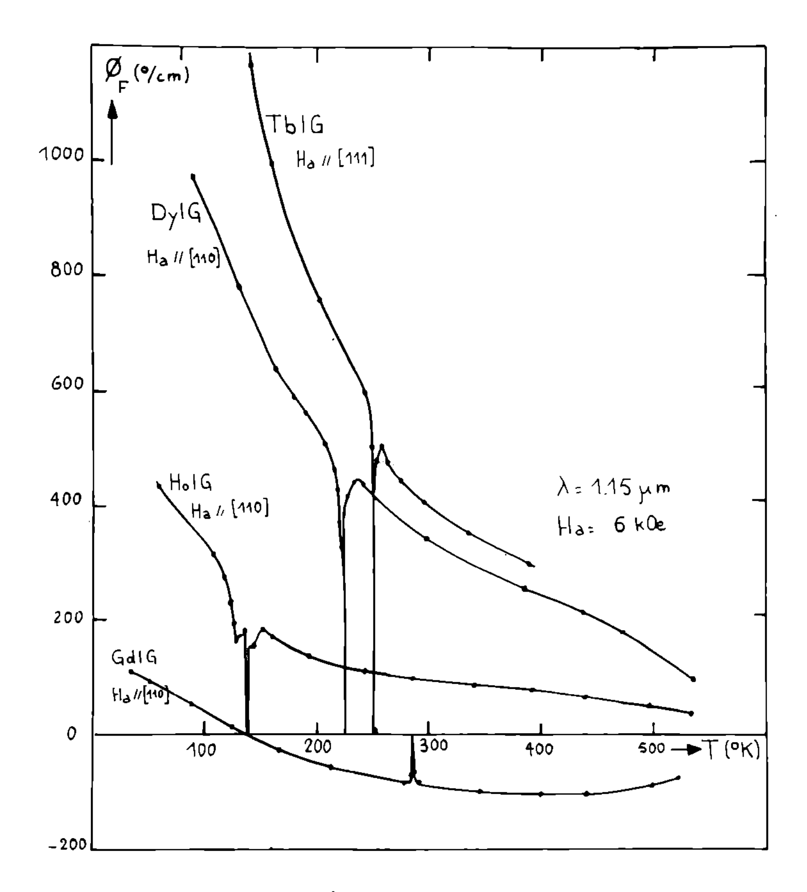


Figure 13

M.O. effect depends on the direction of the magnetization with respect to the crystallographic axis. By using, in cubic crystal, the direction |100|, |010| and |001| as referential, this effect can be described for light propagation along the direction |100| for instance, by the electric dipole P(2)(equation 2) which contains the coefficients f_{12}, f_{44} and Δf.

In most of experimental cases, as Raman scattering and integrated optics, the propagation vector $\vec{q}$ and the magnetization $\vec{M}_0$ are not parallel to the cubic axis, therefore we have to choose a new referential XYZ constructed from the direction either of $\vec{q}$ or $\vec{M}_0$. New M.O. constants f_{IJKL} are deduced from the former $f_{ijk\ell}$ by using the transformation with the Euler angle θ, ϕ, ψ.

When the incident radiation is propagating in the plane 001 with an angle ϕ with respect to the |100| direction as shown in the Figure 15 for instance, the eigen-mode of propagation does not correspond to photons polarized parallel and perpendicular to the magnetization. The eigen modes n_α and n_β are equal to $n_{//}$ and $n_\perp$ only for special values θ

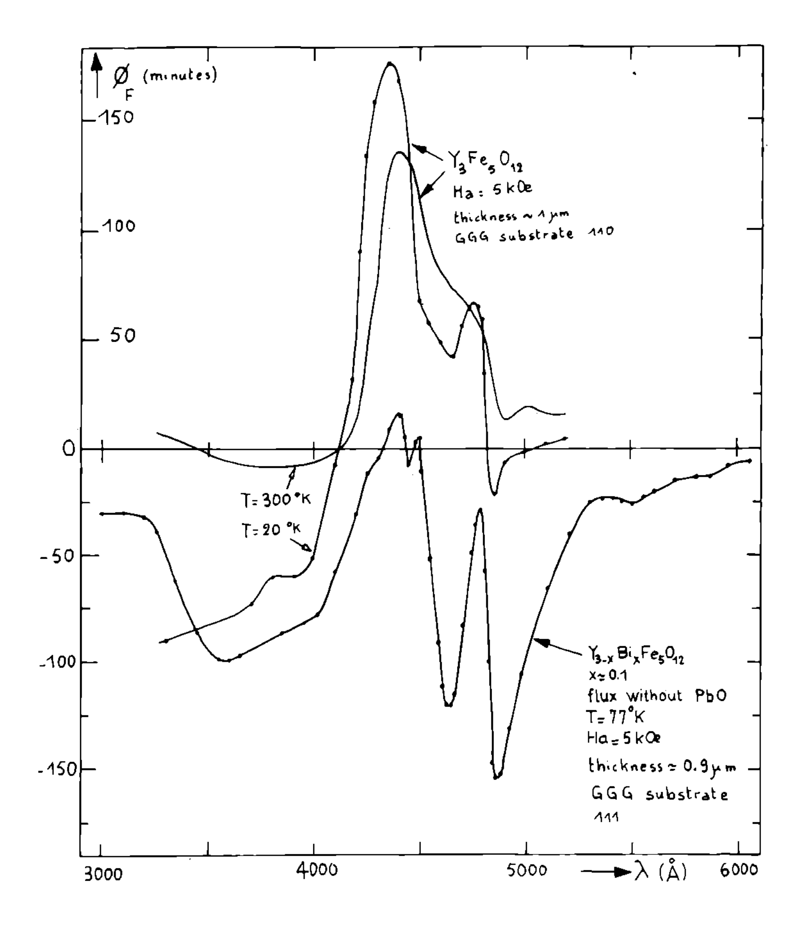


Figure 14

and ϕ as shown in the figure 16. The upper part of this figure gives the LMB for three typical values of the angle θ when $\vec{q}$ is parallel to the |110| direction. The lower part of the figure corresponds to a propagation vector parallel to the |100| direction. It can be shown that the LMB is isotropic only for propagation along one |111| direction.

Typical temperature and magnetic field evolution of ϕ_{CM} in $Y_3Fe_5O_{12}$ and $Tb_3Fe_5O_{12}$ are given in the figure 17 to 20. We note the strong contribution fo the rare-earth ions to ϕ_{CM} mainly at low temperatures, and the large anisotropy of this effect. Similar evolution are observed in most of rare-earth ferrimagnetic garnets.

REFERENCES

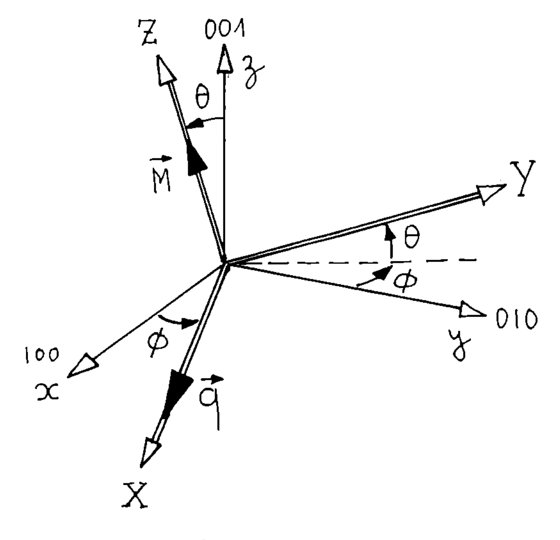
[1] G.S. Krinchik and M.V. Chetkin, Soviet Phys. J.E.T.P. ,14 ,485 (1962).

[2] G. Abulafya and H. Le Gall, Solid State Com., 11 , 629 (1972)

[3] S.H. Wemple , J.F. Dillon, L.G. Van Uitert and W.H. Grodkiewicz, Appl. Phys. Lett., 22 , 331 (1973).

[4] H. Le Gall, A. Lepailler-Malecot, Tran Khanh Vien and M. Artinian, (Internal Report, to be published).

[5] S. Wittekoek and D.E. Lacklison, Phys. Rev. Lett. 28 , 740 (1972).



$$n_\alpha - n_\beta \simeq \frac{M^2}{2\bar{n}}\left\{\left[f_{12}-f_{11}+\Delta f\left(7\sin^2 2\theta+\sin^2 2\theta\cos^2 2\phi+4\sin^4\theta\sin^2 2\phi\right)/8\right]^2+\left(\frac{\Delta f}{2}\right)^2\sin^2 2\theta\left(2\cos 2\theta+\sin^2\theta\sin^2 2\phi\right)\right\}^{1/2}$$

Figure 15

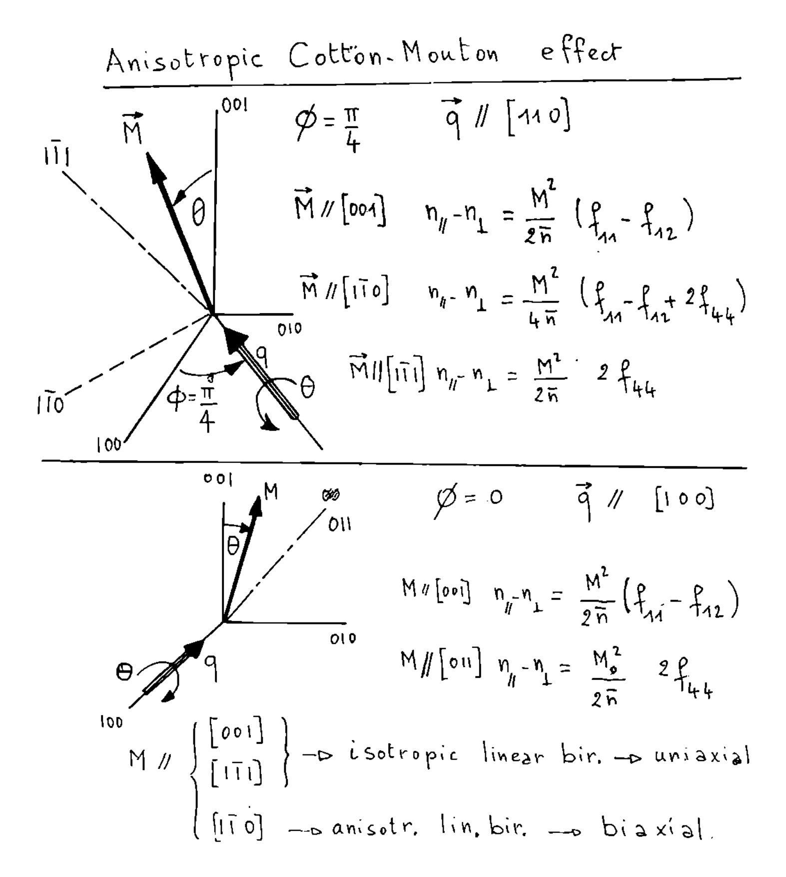


Figure 16

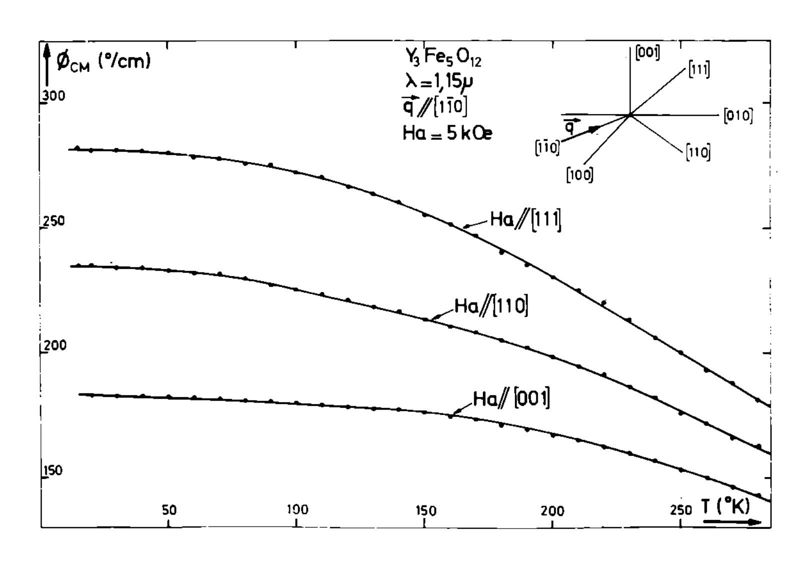
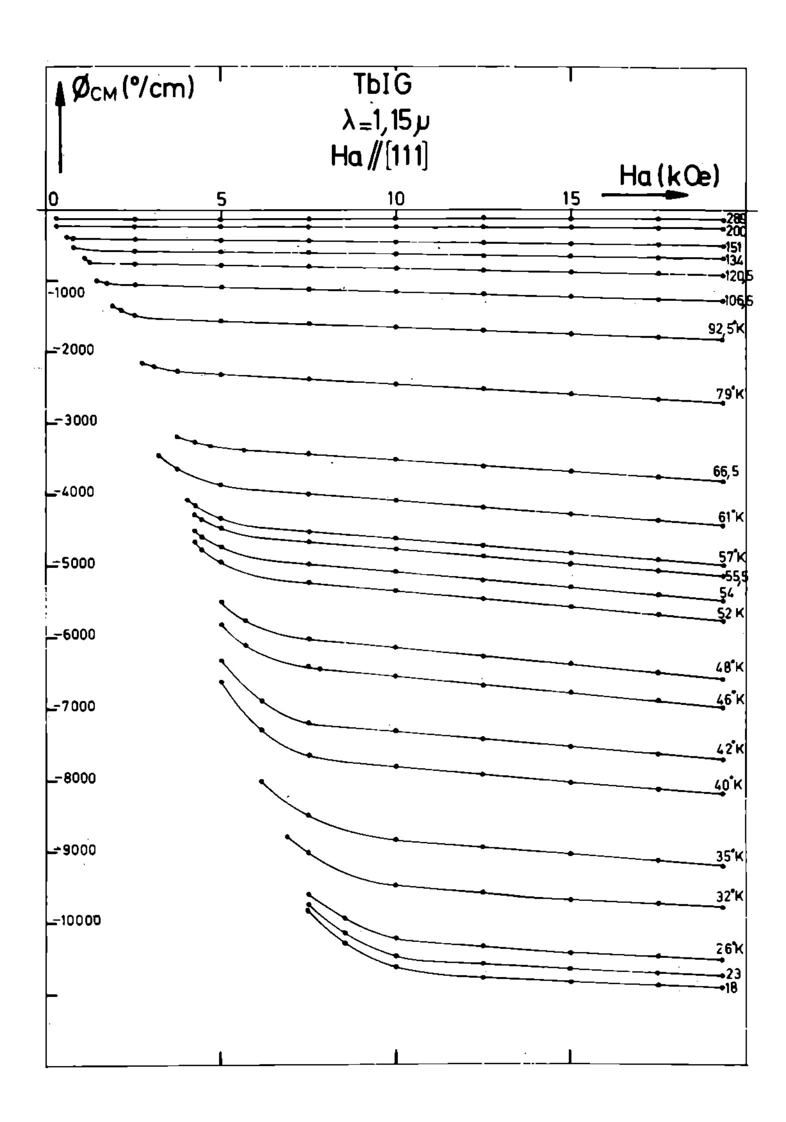
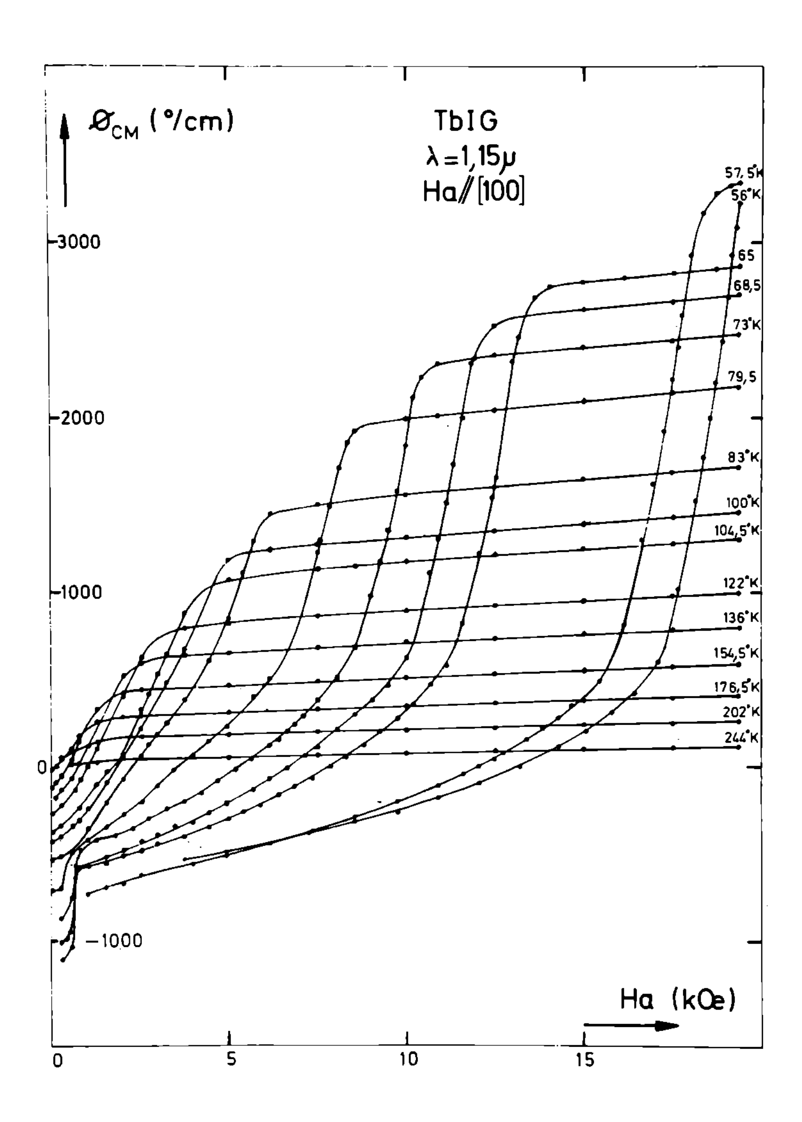


Figure 17



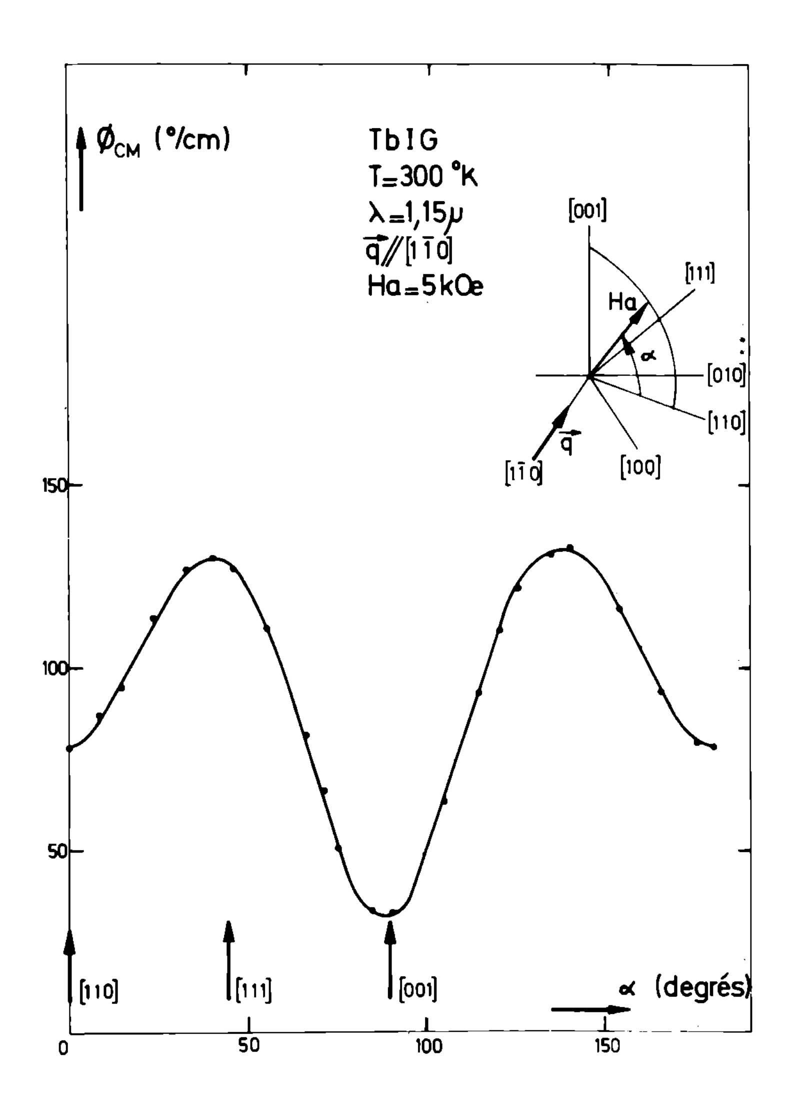
ØCM (°/cm)
TbIG
λ=1,15μ
Ha//[111]
Ha (kOe)
0
5
10
15
-1000
-2000
-3000
-4000
-5000
-6000
-7000
-8000
-9000
-10000
289
200
151
134
120,5
106,6
92,5°K
79°K
66,5
61°K
57°K
55,5
54
52 K
48°K
46°K
42°K
40°K
35°K
32°K
26°K
23
18

Figure 18



ØCM (°/cm)
TbIG
λ=1,15μ
Ha//[100]
57,5K
56°K
65
68,5
73°K
79,5
83°K
100°K
104,5°K
122°K
136°K
154,5°K
176,5°K
202°K
244°K
3000
2000
1000
0
−1000
Ha (kOe)
0
5
10
15

Figure 19



Ø_CM (°/cm)
TbIG
T=300 °K
λ=1,15μ
q // [1 1̄ 0]
Ha=5kOe
[001]
[111]
Ha
α
[010]
[110]
[100]
[1 1̄ 0]
q
150
100
50
[110]
[111]
[001]
α (degrés)
0
50
100
150

Figure 20

BIBLIOGRAPHY ON MAGNETOOPTICS

Theory.

M.J. Freiser, IEEE Trans. on Mag., 4 , 152 (1968).

R. Loudon , Adv. Physics, 17 , 243 (1968).

P.A. Fleury and R. Loudon, Phys. Rev., 166 , 514 (1968)

H. Le Gall, J. de Physique, Suppl. au n° 2-3, 32 ,C1-590 (1971)

Spin-Photon interaction in magnetic crystals Magnetooptical effects : 1 st Summer School on Magnetooptics, Podhradi (1973), Edited by J. Kaczer, Academia Phraha (1974).

Phys. Stat. Solidi (b) , 46 , 467 (1971)

Phys. Stat. Solidi (b), 47 , 591 (1971).

MAGNETOOPTICAL MATERIALS

D.L. Wood, J.P. Remeika, J. Appl. Phys., 38, 1038 (1967)

D.L. Wood, J.P. Remeika, E.D. Kolb , J.A.P. 41 , 5315 (1970)

J.C. Smits, B.E. Ar gyle, M.J. Freiser, J.A.P., 37, 1391 (1966)

F.J. Kahn, P.S. Pershan, J.P. Remeika, Phys. Rev., 186 , 891 (1969)

G.S. Krinchik, M.V. Chetken, Sov. Phys. USPEKHI, 98 , 307 (1969)

V.V. Eremenko, A.I. Belyaeva, Sov. Phys. USPEKHI, 98 , 320 (1969)

G.A. Smolenskii, R.V. Pisarev, I.G. Sini, N.N. Kolpakova, A.G. Titova, Izvestiya Akad, Nauk SSSR, Seriya Fizicheskaya, 36 1219 (1972).

MAGNETO OPTIQUE INTEGREE

B. Desormiere et G. Hepner

Laboratoire Central de Recherche THOMSON-C.S.F.

Domaine de Corbeville - 91401 ORSAY

▪ I. INTRODUCTION

Que signifie tout d'abord ce titre "Magnéto-Optique Intégrée" ? Tout simplement cette partie de l'Optique Intégrée faisant appel aux matériaux magnétiques. Celle-ci, en effet, issue à la fois de l'Optique Cohérente et de la transposition en optique des systèmes à ondes hyperfréquences guidées, requiert la réalisation d'un certain nombre de fonctions élémentaires (émission, détection, modulation, filtrage....). De telles fonctions peuvent être obtenues en utilisant les différents types d'interactions lumière-matière, dont en particulier les intéractions lumière-milieu magnétique.

Les effets magnéto-optiques sont divers (Faraday, Cotton-Mouton, Kerr....) et les matériaux dans lesquels on les observe sont variés (métaux, semi-conducteurs ou isolants, para-, ferro- ou antiferromagnétiques). Cependant, l'Optique Intégrée apporte des contraintes supplémentaires assez sévères (transparence, homogénéité....) limitant le choix des matériaux ayant un intérêt pratique.

Néanmoins, la propagation d'une lumière guidée en milieu magnétique subissant certaines modifications expliquées par des effets magnéto-optiques, a déjà été observée, et ceci avec des caractéristiques suffisamment attractives pour justifier un effort accru des études dans ce domaine.

Dans la première partie de cet article, on s'éfforcera de faire le point des résultats acquis sur le sujet. Dans la deuxième partie, on cherchera à dégager le sens des recherches actuellement en cours.

▪ II. RESULTATS ACQUIS

Lorsqu'on fait la bibliographie des travaux effectués jusqu'à ce jour sur le sujet, on est frappé par la différence entre l'état d'avancement des travaux sur le plan théorique et sur le plan expérimental. L'explication de cette différence réside dans le

fait que si, sur le plan théorique, on peut concevoir différentes structures d'Optique Intégrée utilisant des effets magnéto-optiques, leur réalisation se heurte aux limitations actuelles de la technologie. C'est pourquoi, on passera assez rapidement ici sur des structures qui semblent difficilement réalisables aujourd'hui pour s'étendre davantage sur celles ayant été expérimentées.

A. Les deux types de structure envisageables

Un guide optique est composé d'une couche de quelques microns d'épaisseur déposée sur un substrat. Un milieu supplémentaire peut être rapporté sur cette couche. Il existe donc deux types de structure utilisant des effets magnéto-optiques : ou bien ceux-ci se produisent dans la couche, ou bien dans l'un des deux milieux adjacents à cette couche.

On constate assez curieusement que les premiers travaux théoriques ont eu pour objet le deuxième type de structure alors que les expériences ont jusqu'à présent été conduites sur des structures du premier type. Ceci peut s'expliquer sans doute par le fait qu'il semblait difficilement concevable au début (vers 1970) d'obtenir une couche magnétique de qualité suffisante pour guider la lumière et qu'on pensait pouvoir réaliser plus facilement le dépôt d'une couche de verre, par exemple, sur un substrat magnétique. En fait, stimulée par l'application des mémoires à "bulles magnétiques", l'hétéroépitaxie de Grenats ferrimagnétiques sur un substrat non magnétique a fait des progrès spectaculaires. En fait le substrat est paramagnétique et ses effets magnéto-optiques sont négligeables. Des couches quasi-parfaites optiquement ont pu être élaborées sur des surfaces de plusieurs centimètres carrés avec des homogénéités d'épaisseur meilleure que 0,1 µm. Il se trouve aussi que, par un heureux hasard, l'indice optique de la couche (environ 2,1 à 1,15 µm) est supérieur à celui du substrat (environ 1,9 à 1,15 µm), condition indispensable pour que la lumière puisse être guidée dans la couche. Une telle propagation a été en fait observée, avec des caractéristiques acceptables pour un premier résultat (8 dB/cm à 1,15 µm et 3dB/cm à 1,55 µm)[1]. Nous allons voir maintenant quelles modifications sont prévues par la théorie pour cette propagation de lumière lorsqu'interviennent des effets magneto-optiques.

B. Principe de la conversion de mode par effet magnéto-optique

Deux approches théoriques ont été utilisées pour étudier l'influence des effets magnéto-optiques sur la propagation de la lumière guidée : l'optique des rayons, qui permet de déterminer les modes de propagation du faisceau lumineux en tenant compte des déphasages et des couplages de modes aux interfaces, et l'optique ondulatoire qui permet en principe un calcul complet des modes propres, de leurs indices effectifs, etc.... La première approche est souvent utilisée car elle permet plus facilement de représenter les phénomènes. [2] Elle n'est cependant pas rigoureuse et doit être utilisée avec précaution. La deuxième approche, bien que plus fastidieuse, semble préférable comme première recherche théorique des phénomènes. On sait qu'en propagation libre, une lumière polarisée se propageant en milieu magnétique parallèlement à la direction de l'aimantation subit

un effet Faraday, c'est-à-dire une rotation de la direction de polarisation. On peut donc s'attendre à un effet analogue en propagation guidée. Il existe cependant une différence très importante : la propagation guidée de lumière est régie par des modes, en général non dégénérés. En conséquence, la polarisation ne peut pas être quelconque pour une constante de propagation donnée, comme c'est le cas de la propagation libre. Dans le cas de milieux isotropes et diélectriques, qui est le cas le plus simple, ces modes propres forment deux familles : les modes TE_n (appelés aussi modes H car seul le champ magnétique a une composante longitudinale) et les modes TM_n (ou modes E; l'indice n du mode caractérise le nombre de noeuds dans la variation d'amplitude des champs le long de l'épaisseur du guide). Un exemple typique des courbes de dispersion de ces modes est indiqué sur la figure 1. L'influence des

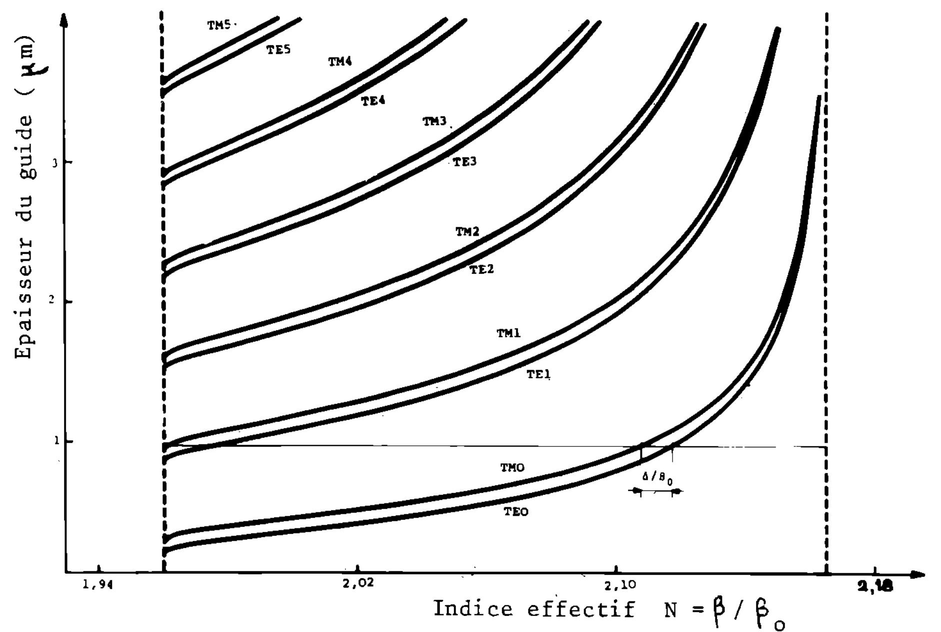


Figure 1 :
Courbes théoriques de dipersion des modes propres pour une couche de Grenat déposée sur Grenat de Gadolinium-Gallium.
β:*constant de propagation* , $\beta_0 = \frac{2\pi}{\lambda_0}$
(λ_0 = *I,I5* µm)

effets magnéto-optiques se traduit dans ce cas par un couplage entre ces modes, comme l'ont établi les premiers S. Wang et al.[2] Du fait de ce couplage, ces modes TE et TM ne sont plus en toute rigueur les modes propres, comme l'ont signalé en premier

S. Yamamoto et al [3]; cependant ce couplage est, en général, suffisamment petit pour qu'une théorie de perturbation puisse s'appliquer. Les modes non perturbés sont les modes TE et TM obtenus en négligeant les effets magnéto-optiques. Cette théorie montre que seuls sont couplés de façon notable les modes TE et TM de même indice. On pourrait dès lors penser qu'il suffit de propager la lumière suivant une distance correspondant à une rotation de 90 degrés pour avoir une conversion complète d'un mode à l'autre. En fait, il n'en est pas ainsi, sans précaution spéciale, du fait de la différence des constantes de propagation (notée Δ sur la figure 1). A. Yariv [4] a montré, dans une théorie générale, qu'un tel couplage entre modes non dégénérés se traduisait par une variation sinusoïdale de l'amplitude des modes qu'on a schématisée sur la figure 2. En particulier, les maxima

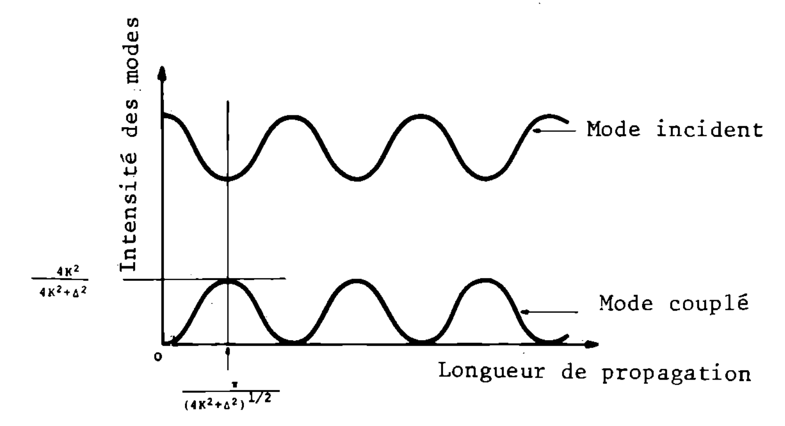


Figure 2 :
Conversion de mode pour une configuration non dégénérée aimantée uniformément.

d'amplitude du mode converti sont typiquement de quelques pourcents par rapport au mode incident, ce qui limiterait beaucoup l'intérêt pratique de cette conversion si elle devait être utilisée telle quelle. Ce n'est heureusement pas le cas car il existe des moyens de surmonter cet inconvénient. L'un de ces moyens, utilisé en premier par P.K. Tien et al, [5] consiste à inverser le sens de l'aimantation le long de la propagation avec une périodicité spatiale adéquate, c'est-à-dire égale à $2\pi/\Delta\beta$, $\Delta\beta$ étant la différence des constantes de propagation des deux modes considérés. Dans ce cas, on obtient un effet cumulatif, le couplage changeant de signe en même temps que le déphasage des deux modes.

C. Caractéristiques des dispositifs expérimentés

Cette conversion de mode peut être modulée en fonction du temps en faisant varier la position de l'aimantation par rapport à la direction de propagation : ceci est obtenu en appliquant un champ magnétique variable. En raison des effets démagnétisants, il est

plus intéressant de faire tourner l'aimantation dans le plan de la couche plutôt que dans un plan perpendiculaire, car les énergies à fournir sont plus petites. Par ailleurs, il est nécessaire d'éviter la formation de parois magnétiques si l'on veut moduler rapidement cette conversion de modes. C'est pourquoi , on cherche à utiliser des couches à "plan facile", c'est-à-dire dont les positions d'équilibre spontané de l'aimantation soient dans le plan de la couche. Or, de telles couches ont pu être fabriquées avec les Grenats ferrimagnétiques en jouant sur les substitutions, de façon similaire, mais en sens opposé, à ce qui a été fait jusqu'à présent pour les couches à "bulles magnétiques" où on cherche à obtenir au contraire un "plan difficile" pour la couche. Le premier type de couche à plan facile réalisée est une couche de composition $Y_3Ga_{1,1}Sc_{0,4}Fe_{3,5}O_{12}$, les substitutions de Gallium et de Scandium permettant d'obtenir une couche épitaxiée (sur $Gd_3Ga_5O_{12}$ orienté (111)) ayant l'axe facile dans le plan, en raison de la désadaptation des paramètres de réseau et des coefficients de dilatation entre la couche et son substrat. De telles couches sont en fait pratiquement à plan facile : le champ d'anisotropie est inférieur à 1 Oe dans le plan et égal à 620 Oe environ suivant la normale à la couche. Des taux de conversion de 52% ont été observés à 1,15 µm en continu. Une profondeur de modulation de 50% a été obtenue à 60 Hz pour un taux de conversion moyen de 15%. La périodicité spatiale de l'aimantation était réalisée à l'aide d'une ligne à méandre placée au-dessus de la couche et parcourue par un courant d'environ 1 Ampère, produisant un champ de 6 Oe environ. Bien que cette valeur soit relativement importante, la très petite impédance du circuit (self de moins de 0,1 µH) en fait un dispositif à faible consommation. La modulation a été observée jusqu'à 80 MHz, limite apparemment due à la réponse en fréquence du photodétecteur plutôt qu'à celle du dispositif. A cette fréquence toutefois, le courant de commande est nettement plus important (12 Ampères environ), ce qui semble pouvoir être attribué à la dynamique de l'aimantation. Néanmoins, la puissance à fournir est d'environ 100 mW, ce qui donne un facteur de mérite du modulateur légèrement supérieur à 1 mW/MHz, valeur tout-à-fait acceptable.

Par ailleurs, ces auteurs ont observé une conversion importante (taux de plus de 40 %) en continu avec un champ appliqué faible (0,2 Oe environ). Ce résultat semble s'expliquer par la présence de domaines magnétiques créant une périodicité spatiale de l'aimantation proche de celle désirée. La difficulté de contrôler et de reproduire une telle configuration limite malheureusement l'intérêt pratique de ce résultat.

Un deuxième type de couche à plan facile a été réalisé et utilisé par des chercheurs d'I.B.M. [6] ; la composition de la couche est $Y_{2,5}Gd_{0,5}Ga_1Fe_4O_{12}$, le substrat étant toujours $Gd_3Ga_5O_{12}$ orienté (111). Dans ce cas, l'existence d'un axe facile dans le plan est obtenue par un autre phénomène, encore relativement mal connu, l'anisotropie de croissance, qu'on peut définir comme un arrangement préférentiel de certains ions lors

du processus d'épitaxie, arrangement différent de celui du cristal parfait. Ce phénomène est cependant bien contrôlé et reproductible. Un taux de conversion atteignant 92% a été observé en continu avec une structure en serpentin parcourue par un courant de 2 Ampères. Un autre moyen de créer une périodicité spatiale de l'aimantation a été experimenté par ces auteurs : il consiste à utiliser des petits rectangles de permalloy disposés au-dessus de la couche le long de la propagation ; leur champ coercitif étant supérieur au champ d'anisotropie dans le plan de la couche, ils jouent le rôle de petits aimants "permanents". Un champ alternatif, d'amplitude suffisante, permet cependant de moduler la conversion de mode en basculant ces aimants. Des taux de conversion de 80% en continu et 70% à 60 Hz ont été mesurés, pour des champs appliqués relativement petits (de l'ordre de l'Oersted).

Enfin, un article rapportant des expériences similaires vient de paraître tout récemment. [7] Des couches de composition $(YLa)_3Ga_{0.33}Fe_{4.57}O_{12}$ ont été utilisées. Le résultat essentiel de cet article est la mise en évidence expérimentale du fait qu'en présence de l'effet magnéto-optique longitudinal les modes propres ne sont plus purement TE ou TM, mais hybrides. Bien que ce phénomène soit peu important pour le matériau étudié (rapport de 3.10^{-3} à 2.10^{-4} environ pour la proportion TE/TM ou l'inverse), il est essentiel de le garder en mémoire, en particulier si on songe à utiliser des matériaux fortement magnéto-optiques.

▪ III. SENS DES RECHERCHES EN COURS

Essayer de dégager le sens des recherches actuellement en cours dans un domaine aussi avancé et aussi évolutif est extrêmement délicat et on doit revendiquer ici le droit à l'erreur. Il est cependant nécessaire de faire une telle démarche pour établir des critères de choix d'orientation des recherches aussi judicieux que possible.

Il est certain que l'aboutissement des recherches en Optique Intégrée soit avant tout la réalisation de systèmes de traitements ou de transmission de données utilisant les fréquences optiques. Néanmoins, il ne faut pas négliger les retombées que peut avoir une telle technologie dans le domaine scientifique, essentiellement en Optique des Solides. Elle conduit en effet à des mesures extrêmement précises de certains paramètres (indices, absorption....) pour des matériaux réalisés en couche fine, là où l'Optique classique atteint ses limites. De même, la grande localisation obtenue dans ce cas pour l'énergie lumineuse devrait donner un regain d'intérêt à des phénomènes tels que l'optique non linéaire ou la microscopie des surfaces.

Il semble prématuré de chercher à savoir quel sera l'impact exact des phénomènes magnéto-optiques dans les systèmes d'Optique Intégrée. En particulier, l'intérêt comparé d'un modulateur magnéto-optique par rapport à celui d'un modulateur électro-optique est actuellement difficile à établir. La réponse dépend en grande partie de la mise au point, qui est encore à faire, de couches adéquates. Ce n'est que très récemment qu'un effort sys-

tématique a été entrepris pour trouver des matériaux fortement magnéto-optiques et en même temps suffisamment transparents. De tels matériaux, s'ils peuvent être obtenus en couches minces d'assez grandes dimensions (1 cm de diamètre environ) doivent accroître de façon importante l'intérêt de ce type de dispositifs (des figures de mérite de l'ordre de 10 µW/MHz ont été annoncées par P.K. Tien au Congrès Intermag 74, pour des Grenats substitués au Bismuth). Par ailleurs, des efforts entrepris pour la mise au point de matériaux transparents à 1,06 µm. Le choix de la longueur d'onde de fonctionnement des systèmes futurs ne peut pas encore être fait, mais cette longueur d'onde risque d'être retenue pour certains d'entre eux, d'où l'intérêt d'une telle recherche. Il faut aussi souligner que l'effet Faraday pour les Grenats ferrimagnétiques, fortement dispersif dans cette zone de longueurs d'onde, croît rapidement lorsqu'on se rapproche du visible.

Un autre point faisant l'objet des recherches en cours dans ce domaine est la compréhension plus détaillée des phénomènes.

Un premier exemple est la détermination exacte des modes propres dans une structure où existent des effets magnéto-optiques. Un tel examen risque d'avoir un intérêt pratique comme par exemple celui de montrer qu'une conversion de mode de 100 % n'est pas possible sans précaution spéciale.

Un autre exemple est l'examen approfondi du comportement dynamique de l'aimantation de ces couches à plan facile lorsqu'elles sont soumises à des excitations magnétiques sinusoïdales ou impulsionnelles. Cet examen, qui fait intervenir des systèmes différentiels non linéaires, est important pour prévoir de façon quantitative la réponse en fréquence de ces dispositifs et l'optimisation des paramètres en fonction des spécifications recherchées. Une détermination plus rigoureuse du facteur de mérite devrait aussi découler des ces calculs.

Un troisième point, qui fait ou fera très probablement l'objet de recherches, est l'interaction d'une lumière guidée soit avec des domaines, soit avec des ondes de spin. Il est difficile de préciser actuellement le type et l'intérêt des recherches à faire dans cette voie, mais il est permis de penser que l'Optique Intégrée est plus compatible que l'Optique classique avec les systèmes qui utilisent ou utiliseront des couches magnétiques, qu'il s'agisse des dispositifs de mémoire ou hyperfréquences.

Enfin, on n'a pas parlé jusqu'à présent du type de dispositifs qui est spécifique aux matériaux magnétiques, les dispositifs non réciproques. Par comparaison avec ce qui existe actuellement en hyperfréquence, on peut penser que de tels composants, isolateurs et circulateurs, joueront un rôle important dans les systèmes d'Optique Intégrée. Or, si plusieurs structures ont été proposées théoriquement pour réaliser de tels dispositifs et ceci dès le début des travaux effectués en Magnéto Optique Intégrée, aucune n'a été effectivement réalisée à ce jour, du moins à notre connaissance. La raison semble être la difficulté de réaliser une structure composite permettant d'obtenir une conversion de mode non réciproque, obtenue par les effets magnéto-optiques, plus une conversion réciproque, obtenue par l'emploi d'un matériau biré-

fringent. La proposition faite par J. Warner, [8] de rapporter par contact optique sur la couche un matériau biréfringent convenablement orienté ne semble pas une solution très réaliste dans une perspective d'industrialisation et la bonne solution, utilisant un matériau biréfringent intégré à la structure, reste à trouver.

IV. CONCLUSION

L'importance future de la magnéto-optique intégrée, très jeune partie (3 ans) de ce jeune domaine (6 ans) qu'est l'Optique Intégrée est difficile à saisir. En se référant aux progrès accomplis depuis son début, où l'on ne pensait pas, semble-t-il, pouvoir un jour utiliser des couches magnétiques comme guides de lumière, il est permis de penser que ce domaine va évoluer rapidement dans des directions qu'il est difficile de préciser avec certitude. Cette évolution sera d'autant plus rapide qu'elle est motivée par la mise au point de dispositifs spécifiés que sont les modulateurs et les composants non réciproques, ces dispositifs permettant de mieux préciser les travaux restant à faire, à la fois dans le domaine théorique et dans celui de la technologie. En particulier les prochains mois ou les prochaines années devraient être décisives pour démontrer la possibilité de réalisation de composants non réciproques.

REFERENCES

[1] P.K. Tien, R.J. Martin, S.L. Blank, S.H. Wemple and L.J. Varnerin Appl. Phys. Lett. ,21 ,5, pp. 207-209 1er Sept. 1972

[2] S. Wang , J. Grow and M. Shah
I.E.E.E. Trans on Mag. V- Mag 7, pp. 385-387 Sept. 1971
Appl. Phys. Lett., 19,6, pp. 187-189 15 Sept. 1971
I.E.E.E. Journal on Quantum Electronics, V, QE-8, 2, pp. 212-216 Fév. 1972

[3] S. Yamamoto, Y. Koyamada and T. Makimoto
J. Appl. Phys. , 43 , 12 , pp. 5090-5097 Décembre 1972.

[4] A. Yariv I.E.E. Journ. of Quant. El., QE-9, pp. 919-933 Sept. 1973

[5] P.K. Tien, R.J. Martin, R. Wolfe, R.C. Le Cram and S.L. Blank Appl. Phys. Lett., 21 ,8 pp. 394-396 15 Octobre 1972.

[6] S.C. Tseng, A.R. Reisinger, E.A. Giess and G.G. Powell Appl. Phys. Lett. 24 , 6, pp. 265-267 15 Nov. 1974.

[7] J. Warner Appl. Optics , 13 , 5, pp. 1001-1004 May 1974

[8] J. Warner I.E.E.E. Trans. on MTT-21 , 12, pp. 769-775 Décembre 1973.

HOLOGRAPHIC STORAGE

A.M. Glass

Bell Laboratories

Murray Hill, New Jersey 07974 U.S.A.

I. INTRODUCTION

The potential of high capacity information storage using optical techniques has been realized for many years. Long before the discovery of the laser photochromics were investigated for possible application to high resolution erasable optical memories as a possible alternative to permanent, granular photographic techniques. With the advent of the laser came the capability of addressing extremely small regions of a recording medium, of dimensions limited only by the diffraction of light, thereby allowing information storage densities in excess of 10^{8} bits/cm^2 . Optical holography promised to even further increase this capability by utilizing the thickness dimension of the recording medium to achieve bit densities $\sim 10^{12}$bits/cm^2 . Furthermore rapid random access of this information should be achievable without serious interconnection problems. When viewed in the context of present day magnetic storage of bit density $\sim 10^5$ bits/cm^2 and the necessary trade off between capacity and random access time due to the mechanical motion of tapes or discs, the holographic techniques seem an inevitable advance.

Although practical considerations such as the optical arrangement and signal to noise do not allow the utilization of the full capacity potential of the storage media, significant advances have been made toward the development of optical mass storage system. A 10^{13}bit non-holographic optical memory [1] , based on strip recording, has already resulted in ten-fold faster random access times than magnetic tape, lower error rate and considerable space savings due to the increased storage density. The feasibility of holographic memories with gigabit/second bit rates both during recording and sequential retrieval of information have been demons-

trated with a moving holographic recording plane [2,3]. Holographic memories have been built without mechanical motion of the memory plane, each hologram randomly accessed on a microsecond time scale [4,5,6]. One of these is a compact commercial read only memory with a total capacity of 56×10^6 bits (100 cm^2 storage area) with a random access time of 10 μsecs. Only one experimental system [6] has used the third dimension afforded by volume holography, primarily because at the present time no recording medium is completely satisfactory for volume holographic storage. Neither is there a material adequate for a practical in-situ read-write-erase system. The emphasis of this article will be placed on the requirements for the storage medium and the hopes and limitations of the materials which have been investigated in recent years. To put these requirements into perspective the following two sections will be devoted to the underlying principles of holographic recording and a brief discussion of the entire memory systems which are at present being proposed.

II. DIFFRACTION EFFICIENCY AND RECORDING SENSITIVITY

All optical recording media depend firstly on the absorption of radiation containing the information to be recorded and secondly on the change of the absorption coefficient α, or the refractive index n at the wavelength of the read beam, or of the thickness of the medium. An elementary hologram of spatial frequency K is formed by the interference of two plane waves at an angle 2θ as shown in Figure 1, where $K = 4\pi \sin\theta / \lambda$, λ being the wavelength of the recording radiation. The reconstruction efficiency η is defined as the ratio of the diffracted intensity

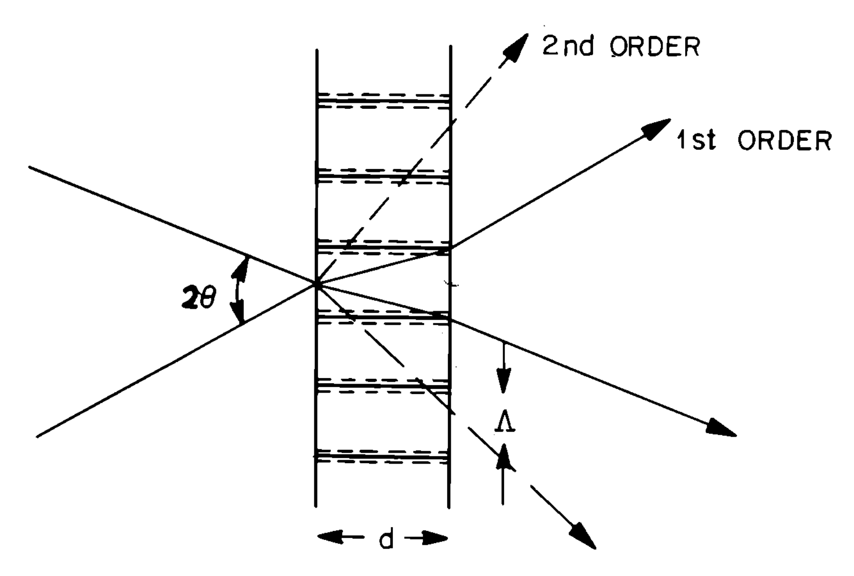


Figure 1 :
Schematic of the formation of an elementary hologram of fringe spacing Λ in a material of thickness d.

contributing to the reconstruction of the original wavefront to the intensity of the incoming read beam. For thin transmission holograms where the recording depends on the change of transmission ΔT (peak to peak) due to a change of α, the diffraction efficiency [7]

$$\eta_a = \frac{(\Delta T)^2}{16} \quad (1)$$

For thin transmission holograms based on the change of index or thickness

$$\eta_p = e^{-\alpha d/\cos\theta} \; J_1^2 \left(\frac{\pi \Delta nd}{\lambda \cos\theta}\right) \quad (2)$$

where Δnd is the peak to peak change of the product nd. J_1 is a Bessel function of first order.

The maximum values of η_a and η_p for thin holograms are 6.25 % and 33.9 % . Part of the energy goes into other diffracted orders.

A hologram is considered thin if d is much less than the grating spacing $\Lambda = 2\pi/K$ of the elementary hologram. There is no wavelength selectivity or angular selectivity of thin holograms so they can be reconstructed from any angle and with any wavelength radiation, although the diffraction angle will vary in either case. In a thin recording medium which requires no development (such as photochromics) the holograms form in real time so that the different diffracted orders and the undiffracted beams can interfere during recording, further reducing the reconstruction efficiency.

For thick transmission holograms based on a peak to peak change of absorption coefficient $\Delta\alpha$ [7]

$$\eta_a = e^{-\frac{\alpha d}{\cos\theta}} \; \sinh^2 \frac{\Delta \alpha d}{4\cos\theta} \quad (3)$$

and for thick phase holograms

$$\eta_p = e^{-\frac{\alpha d}{\cos\theta}} \; \sin^2 \frac{\pi \Delta nd}{2\lambda\cos\theta} \quad (4)$$

For Bragg diffraction from thick phase holograms reconstruction is only possible within a limited range of angles of incidence of the reconstruction beam and a limited wavelength range. Only under certain circumstances(such as elementary holograms) can the wavelength of the readout beam be varied by changing the angle of incidence.
For phase hologram the angular selectivity is

$$\Delta\theta \sim \frac{\Lambda}{d} \quad (5)$$

and the wavelength selectivity

$$\Delta\lambda \sim \frac{\Lambda \cot \theta}{d} \tag{6}$$

For instance for a medium of 1 cm thickness and a spatial frequency of 1000 lines/mm , $\Delta\theta \sim 0.1$ mR and $\Delta\lambda \sim 2$Å , at 5000 Å . For faithful reconstruction of a thick hologram containing several spatial frequencies readout with the same wavelength and angle used for writing is essential. While this poses some problems of alignment of the hologram for readout it has the advantage that individual holograms can be superimposed in the same volume of material either by varying the wavelength or angle of the writing beam outside the ranges given in Equations 5 and 6. Up to 1000 holograms can in principle be superimposed with an angular variation of only 0.1R in 1 cm thick media.

The maximum value of η_a in Equation 3 is 3.7 % while the maximum value of η_p can approach 100 %. This value of η_p can only be obtained for weakly absorbing media and corresponding low recording sensitivity. For most systems the index change Δn is proportional to the absorbed energy density so that at a distance x below the illuminated surface.

$$\Delta n \propto W_0 \, \alpha \, e^{-\alpha x / \cos\theta} \tag{7}$$

for incident energy W_0 so that the maximum recording sensitivity $\Delta n / W_0$, obtained from Equations 4 and 7 is achieved for a crystal transmission of 1/3 with a corresponding maximum diffraction efficiency of 33 %. However, recording by multi-photon absorption (see section XIII) allows 100 % absorption of the recording radiation, yet with reconstruction efficiencies of 100 %.

A useful characterization of the recording sensitivity of a material is [8]

$$S = \frac{\sqrt{\eta}}{W_0 V}$$

where V is the fringe visibility of the incident interference pattern. For many media the sensitivity can be varied at will by varying the optical absorption of the active centers, then a useful characterization is $S' = \sqrt{\eta} / \alpha W_0$ for V = 1. This is the diffraction efficiency per absorbed energy.

▪ III. HOLOGRAPHIC MEMORY SYSTEM

Although the principles of holography have been applied successfully to sequentially addressed memories utilizing mechanical motion of the memory plane, with very high bit rates, the primary concern of this article is in randomly accessed, inertialess,

page organized holographic memories which offer the potential of rapid random information retrieval. For such a system holographic recording has the advantage that, since each bit of information is contained in the entire hologram (at the Fourier transform plane) the spatial redundancy relaxes the requirement for accurate positioning and high optical quality necessary for micro-image and bit organized systems. Furthermore, with suitable recording and readout techniques the hologram can act as its own optical imaging system. The holographic memory also appears well suited for parallel associated search for information [9,10].

Typical elements of a read/write holographic storage system are shown in Figure 2 [11]. In this arrangement the same laser is used for both storage and retrieval of the information. The deflector and optical arrangement displaces the beam parallel to itself in the X-Y plane.

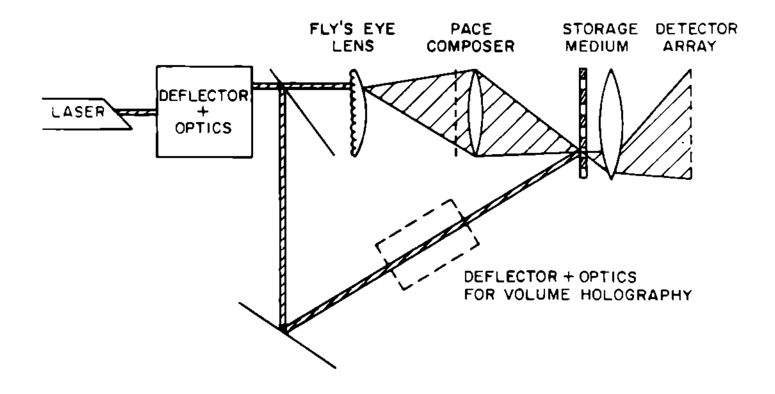


Figure 2 :
A basic arrangement for a holographic read-write memory. [11]

For writing, the beam is split into an object beam and a reference beam, the latter being directed to a particular element of the storage plane determined by the deflector. The object beam is defocussed by an element of the fly's eye lens to illuminate the page composer which is an X-Y array of light valves, each valve containing a single bit of information to be recorded. An object lens is placed close to the page composer so that the Fourier transform of the light valve array is produced at the storage plane at the same point as the reference beam. The resulting interference pattern of the two beams is recorded in the storage medium in the form of a hologram.

For reading, the reference beam only is diffracted from the hologram and an image of the page composer is projected through a field lens onto the photodetector array. The signal from each

detector corresponds to one bit of the page composer. Deflection of the laser beam to a new position allows recording (or reading) of a different portion of the memory plane, yet for each position the entire page composer is illuminated and automatically imaged on the detector array. The deflector and optical system shown in the reference arm of Figure 2 is to enable changing the angle of incidence of the reference beam at the recording plane to allow superposition of holograms in a thick recording medium.

Several variations of this basic system have been proposed and utilized, so the basic approaches will be briefly summarized. The laser and deflector systems may be replaced by a laser array. A read only holographic memory is commercially available which utilizes a thin film CdS laser end pumped with an electron beam in a cathode ray tube configuration. This device produces about 10^6 individually addressable spots of 25 μm diameter, each of which addresses a single page of the memory with random access times of 10 μsec. The laser energy and coherence is too low for holographic recording. Similarly injection lasers have too low an energy for holography but they have been successfully used for bit organized memories [12,13]

At the present time a single laser looks the most promising approach for in-situ writing. Lasers operating on a cw basis or repetitively pulsed with adequate coherence are readily available. Optical deflectors currently available satisfy the requirements for rapid access to different pages of the memory plane. These have recently been reviewed by Zook [14]. With all deflectors, whether mechanical, acousto-optic or electro-optic there is a trade off between access time and the number of resolvable spots. The present performance of acousto-optic devices for an access time of about 2 μ seconds is about 500 resolvable spots, limited by the transducer. Thus two defectors (X and Y) in series can give random access to over 10^5 holograms on a microsecond time scale.

The page composer of Figure 2, containing the information to be recorded in a single hologram is usually considered to be an electrically or optically addressed array of light valves, although for some applications it could be an analogue transparency. Materials under consideration for the page composer are nematic liquid crystals [4,6,15], lanthanum doped lead zirconate titanate ceramics (PLZT) [16], ferroelectric-photoconductor layers (see section X) [17], ferroelastic Gadolinium Molybdate, [18] thin CdS films (depending on the temperature shift of the absorption edge) and acousto-optic modulator arrays [2,3]. In the acousto-optic device the information is carried by several acoustic beams generated in parallel by an array of transducers. A primary requirement of the page composer is that the light valve on/off contrast ratio should be high to maximize signal/noise in the hologram. For many applications the data composition rate should be comparable to the data access rate of the memory, although for read mostly applications a slower page composer backed up with a small buffer memory can provide adequate

system performance.

Currently available integrated silicon detectors have adequate sensitivity to cope with the demands and capacity of most of the proposed memories. For instance a system utilizing a 1 watt laser, 10^4 bits per page and an access time of 1 μsecs and an overall system efficiency of 1% requires a detection capability of 10^{-12} Joules. This figure can be met with adequate signal/noise with existing technology. Charge coupled devices are capable of detecting 10^{-13} Joules with a signal to noise ratio of 23 db with data transfer rates > 10 MHz.

Several variations of the optical arrangement have been proposed. For instance, the fly's eye lens can be replaced with an array of holograms in dichromated gelatin which serves the same purpose. The hologram array can be recorded in such a way as to eliminate the field lens [19] .Using reflection holograms instead of transmission holograms, the page composer and detector array have been combined into a single compact unit [20] .

The total capacity of such a memory scheme is the product of the number of resolvable spots of the main deflection system, the number of bits per page and the number of superimposed pages in each region of the recording medium. From detailed considerations [19,21] of the diffraction from the limiting apertures of the optical system, of the spacings between detectors in the array and of holograms in the memory plane for adequate signal to noise and other geometrical considerations the estimated total capacity is limited to $\sim 10^8$ to 10^9 bits without superposition of pages. This may be increased to $\sim 10^{10}$ to 10^{11}bits by superposition of pages in thick recording media. It is clear that this limitation is not set by the recording medium but by practical considerations of the optical arrangement. In a real system some further reduction in capacity may be necessary to decrease the error rate [4] in a high capacity system.

▪ IV. STACKED OPTICAL MEMORIES

Two schemes have been proposed for increasing the capacity of a holographic memory without increasing the access time by stacking several holograms in series. This differs from thick holograms in that for a given interrogation angle the hologram from any individual stack may be read separately. The first scheme [22], shown in Figure 3a uses media in which the holographic recording and diffraction efficiency can be controlled with an applied field. Such materials are discussed in section XII. Holograms can only be recorded or read in that hologram plane in the stack which has an externally applied field, all other members of the stack being inactive. Thus any hologram in the stack may be addressed by the appropriate combination of beam position and electric field.

Such a scheme implies that the recording medium is reasonably transparent since the read beam has to pass through the entire stack. This in turn implies some sacrifice of recording sensitivity unless multiphoton absorption is used (see section XIII)

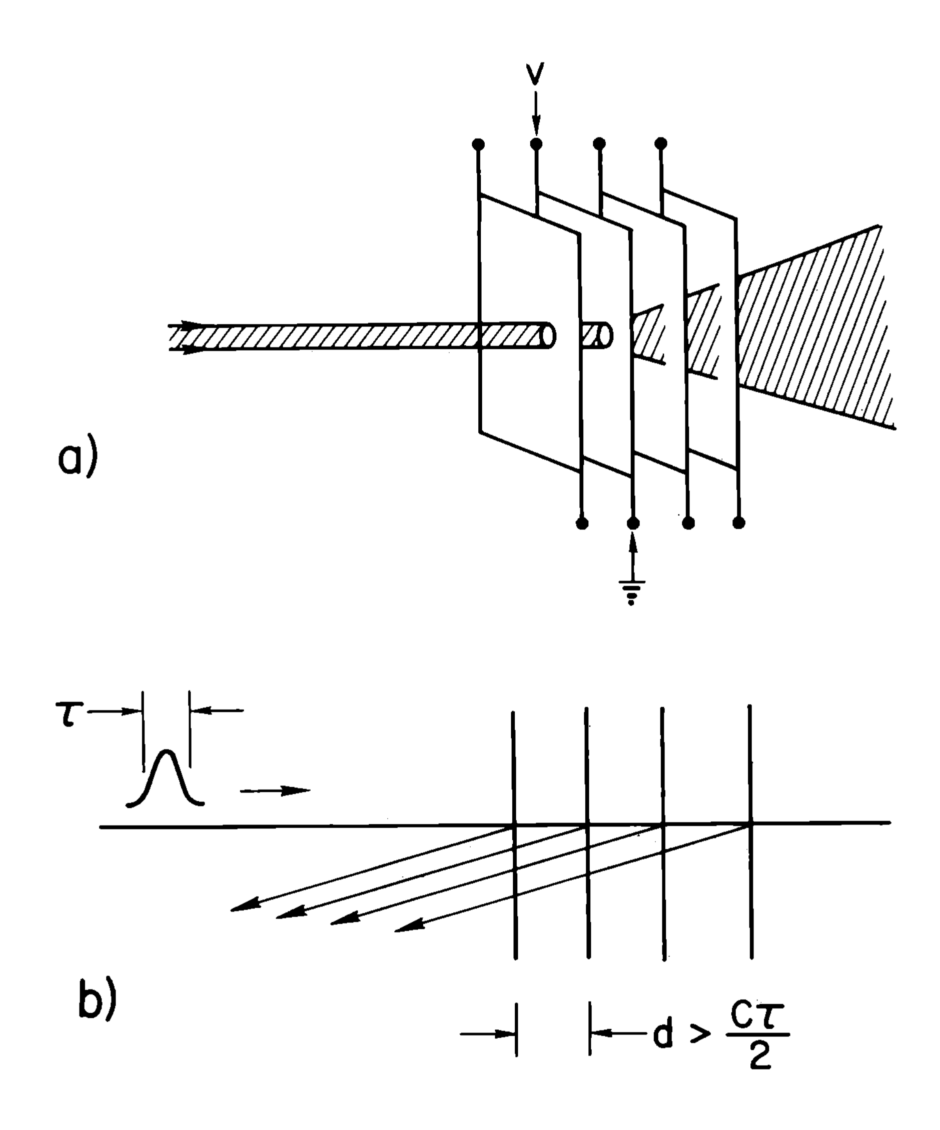


Figure 3 :
Stacked holograms. a) using electrically sensitized media b) using time of flight separation of holograms.

A second proposal [23] uses the time of flight of picosecond optical pulses to separate the image reflected from individual recording planes in the stack as shown in Figure 3b. The planes must be separated by more than $\frac{1}{2}$ cτ where c is the light

velocity and τ the pulse length. Such a scheme is wrought with difficulties. To avoid a bulky stack very short pulses are necessary making in situ recording difficult because of the short coherence length of such pulses, and making detection very difficult since the detector must be fast and sensitive enough to resolve the pulses. Furthermore, the recording medium must be reasonably transparent, and the readout process is necessarily noisy since the reconstructed beam from the back of the stack is scattered from the successive storage planes.

V. THE IDEAL STORAGE MEDIUM

The ideal storage medium for an optical memory would have the following attributes :

a) Read-write-erase in situ. b) High resolution. c) High recording sensitivity. d) No degradation during readout. e) Long storage time. f) High signal/noise. g) Recording sensitivity independent of spatial frequency. h) Writing at a convenient laser wavelength. i) Linear recording characteristic (see below). j) Room temperature operation. Most of these requirements are expressed qualitatively since any improvement is desirable, but to be somewhat more specific preferred values are : b) $\geq$2000 lines/mm, so that bit density is diffraction limited, c) <1 mJ/mm^2 . This value is set by cw laser powers of 1 watt and recording times $\sim$ 1 msec., e) This depends on the application. For archival storage many years is necessary, while for applications involving a periodically revised data block the storage period could be much shorter. f) $\geq$ 20 db, corresponding to a theoretical error rate of 10^{-11} [24] , i) linear above a well defined threshold may be desirable to provide some immunity from low level scattered light, j) within range of convenient heating or cooling devices.

At present no material satisfies all these requirements, and a considerable amount of materials research is still necessary. However, when a material approaching the ideal is discovered it will find extensive application. In the following sections the various materials which have been investigated in recent years, and the manner in which they satisfy or fail the above requirements are discussed. A brief discussion of silver halide emulsions is included since these still dominate the field of optical information storage and represent a goal for other recording materials for sensitivity and stability. Since an understanding of the basic storage mechanisms is necessary for full evaluation of the storage techniques and for improvements a description of the fundamental principles is presented to a greater or lesser degree depending both on the familiarity of the process and on the present degree of understanding.

VI. SILVER HALIDE EMULSIONS, PHOTORESISTS

Photographic materials already meet the requirements of present

read only holographic systems and future improvements will probably await the requirements of a commercial system. The resolution of commercially available emulsions exceeds 2000 lines/mm with exposure requirements of 100 μJ/cm^2 for Kodak 649 F (wide spectral sensitivity) and 1 to 10 μJ/cm^2 for emulsions of narrow spectral sensitivity such as Agfa 10 E 70 , for a few percent diffraction efficiency. The maximum diffraction efficiency of such an absorption hologram is about 6% but this can be increased by bleaching the hologram. However, bleaching decreases the signal to noise ratio and higher exposures (∿500 μJ/cm^2) are required. For some applications shrinkage of the emulsions during development can be a problem. An important step toward obtaining high quality, high diffraction efficiency holograms for read only systems was the use of hardened dichromated gelatin. [25] With this material the thickness is great enough to get the high diffraction efficiency of volume holograms, but not so great as to get narrow angular aperture. Increased exposures of about 20 mJ/cm^2 are required.

There has been some recent interest in photoresists, which form imaged relief patterns upon exposure and development, due to their advantage in mass replication[26]. Replicas of relief holograms can be embossed into transparent vinyl, the resulting phase holograms being due to modulation of the thickness. Photoresists are organic materials which form crosslinks in the illuminated region (negative photoresist) or else crosslinking is removed by the radiation (positive photoresist). The development process dissolves away the soft regions containing no crosslinks.

An extensive study of the commercially available Shipley AZ 1350 photoresist was recently undertaken by Bartolini [26] . He found that resolution in excess of 1500 lines/mm could be obtained with a signal to noise ratio greater than 25 db. Diffraction efficiencies of 5% were obtained with an exposure of a few millijoules/cm^2 .

A dry processed photoresist has recently been reported [27] in which the surface relief variation is due to shrinkage which is attributed to the polymerization of a highly reactive monomer in the presence of a halomethane photoactivator. The activator inhibits polymerization in the dark region. Following illumination the film is heat processed at 160°C which volatilizes the unreacted activator from the dark region which allows the unreacted monomer to polymerize. This treatment greatly enhances the surface relief. The procedure produced holograms with a maximum diffraction efficiency of 30% with an exposure of ∿ 2 mJ/cm^2 at a spatial frequency of 250 lines/mm. At higher spatial frequencies the sensitivity decreased rapidly to ∿4% for 8 mJ/cm^2 at 1000 lines/mm. While this is a serious limitation the photoresist may find application for in situ recording.

Although the sensitivity of these materials is high the limitations are clear. They have no possibility of erasure and the thickness of the films does not allow superposition of holograms. The goal for further materials research is to overcome these major obstacles.

VII. THERMOMAGNETIC MATERIALS

Optical storage in magnetic materials is based on the Faraday rotation of plane polarized light transmitted along the axis of magnetization. The sense of the rotation depends on the direction of the magnetization M_s. Local changes of M_s are produced by heating a thin film of the material with intense light pulses to a temperature at which the magnetization can be reversed, in the illuminated region only, with an external magnetic field. To erase the information the regions are reheated with light pulses with the magnetic field applied so as to restore the initial direction of magnetization.

For such a storage system magnetic materials are required with the easy direction of magnetization normal to the memory plane, with a very high coercivity at the ambient temperature (preferably room temperature), but with a temperature region of low coercivity not too far removed from this temperature. Furthermore, since thin films are required to keep the thermal capacity of the films small, materials of high specific Faraday rotation F and low optical absorption α at the readout wavelength are required. Of course at the recording wavelength high absorption is necessary to achieve sufficiently high temperature change. The magneto-optic figure of merit for readout is usually expressed as $2F/\alpha$. Unfortunately at wavelengths where α is small F is small also.

The films which have been most extensively investigated are Manganese Bismuth films [28] for which F is $\sim 7.5 \times 10^5$ degrees/cm and $2F/\alpha \sim 3.5$ degrees at 6328 Å. While these qualifications are favorable for readout, the films have to be heated to the Curie temperature at 360°C (in the normal phase) for magnetization reversal requiring high optical intensities for writing. For Curie point writing an external magnetic field is not essential since the surrounding magnetic field forces a closure in the heated spot when it cools through the Curie temperature. However, an external field improves the recording sensitivity.

An advantage of these films is that the information may be stored for long periods of time without degradation and they show no fatigue with repeated cycling of the information.

Because of the high intensities needed for writing almost all the investigations of magneto-optic materials have involved bit by bit organized memories for which the high intensities can be readily achieved with lasers of moderate power by focussing to diffraction limited spots 1 to 10 μm diameter. However Mezrich [29] has recorded elementary holograms in Mn-Bi films of about 1000 lines/mm resolution using pulses from a Q switched ruby laser 15 nseconds in duration and about 1 MW/cm^2, resulting in a readout efficiency η up to 5×10^{-4}. Although this efficiency is low improved signal to noise is obtained with polarizers since the diffracted wave is polarized at 90° to the incident wave. These experiments showed that recording in Mn-Bi films is not just a simple threshold process but that there is

some limited linearity of photoresponse.

The energy density required for writing using a thermal process of this kind can be easily estimated from the specific heat c_v and the film thickness d. i.e

$$W_0 \sim c_v \, d\Delta T \qquad (9)$$

assuming 100 % absorption of the radiation. ΔT is the necessary temperature rise. Since c_v is typically 10 Joules/cm°C for these materials (at room temperature), for a film thickness of 10^{-5} cm we have for Mn-Bi the minimum required energy density.

$$W_0 \gtrsim 3 \times 10^{-2} \text{ Joules/cm}^2$$

For high bit densities this energy must be delivered in a time short compared with the thermal diffusion time to surrounding regions. Mezrich[29] has shown that for a holographic fringe spacing Λ the optical pulse duration τ must satisfy.

$$\tau \ll \frac{\Lambda^2}{4\pi^2 k} \qquad (10)$$

where k is the thermal diffusivity. Thus for a resolution of 1000 lines/mm a pulse duration $\tau < 10$ n secs is required. Any increase of τ will result in decreased fringe contrast. The required peak intensity is thus about 1 MW/cm^2 as found experimentally. Lasers of sufficiently high pulse intensity and energy have very slow repetition rates which at present would limit the usefulness of such a system.

To improve the sensitivity only two approaches can be taken. Firstly suitable films could be operated at a low temperature where c_v is greatly reduced. Increased writing sensitivity has been obtained in this way with EuO [30], which has a low Curie temperature, operated at 25°K. However, operation at cryogenic temperatures is not likely to achieve wide application. Secondly, materials with a Curie temperature close to room temperature, or a rapid temperature dependence of the coercivity can be used to reduce the temperature change which must be provided by the light beam. In this latter category Ferrimagnetic materials are interesting since the coercivity is extremely high at the compensation temperature (where the magnetization of the two sublattices are equal) dropping rapidly just a few degrees above this temperature. The compensation temperature of Gadolinium Iron Garnet [31, 33] is around 15 °C and a temperature rise of only 3°C is necessary for magnetization reversal. The compensation temperature of Gd-Co can be "tuned" over a wide temperature range by varying the composition [34] . For a 3° C temperature change, which seems to be the smallest practical value, the incident energy density calculated from Equation 9 must satisfy

$$W_0 \gtrsim 10^{-3} \text{ J/cm}^2$$

for room temperature operation, or intensities greater than 100 KW/cm^2 - still too high to be practical.

The large number of magnetic materials which have been investigated for bit organized memories have been tabulated by Chen[35].

VIII. AMORPHOUS SEMICONDUCTORS

Holograms have been recorded in several amorphous semiconductors [36,38] with light of energy greater than the electronic bandgap of the material. With As-S films [36] of various compositions, holograms of reconstruction efficiency (at 6328 Å) up to 18 % were recorded with an exposure of about 10 J/cm^2 at 4880 Å and a film thickness of 0.045 cm. The illumination was accompanied by a shift of the bandgap to longer wavelengths and an associated change of refractive index. Optically induced index changes up to 0.056 [37], diffraction efficiencies up to 80 % at 6328 Å [38] and resolution up to 2860 lines/mm [38] have been demonstrated in As-S compounds.

Hologram formation has been attributed to crystallization of the amorphous film when it is heated above the glass transition temperature, but below the melting point, by the light. [39] The presence of free carriers appears to enhance the rate of crystallization allowing recording on a microsecond time scale [39]. However this explanation does not account for all the observations since holograms have been recorded over long periods of time with low intensities which do not heat the film above the glass transition, so it appears that the free carriers alone effect some reaction. [38]

For erasure the film may be returned to its initial state by heating it above the melting point followed by rapid quenching.

The reverse processes have also been used for recording and erasing in Te-Ge-As films. i.e. starting with the films in the crystalline state and recording by heating the film above the melting point [40], and erasing by heating between the glass transition and melting point.

In addition to the amorphous crystalline transition it seems that other reactions are occurring under the influence of light. Enhanced precipitation of sulphur microcrystals [41] in As_2S_3 suggests that the reaction [42].

$$2\,As_2S_3 = 2S + As_4S_4$$

may be occurring in the presence of light. Precipitation of Te crystals in Te-Ge-As films [40] were observed during crystallization and they disappeared when the film was returned to the amorphous state.

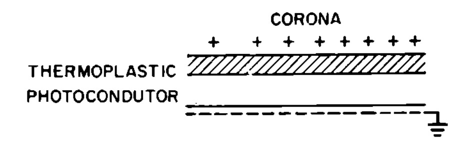
Long term storage of holograms in amorphous semiconductors seems feasible. No deterioration of information stored in As-S

films was observed 2500 after hours. However, the low sensitivity of these films seems to be a limitation for holographic recording, although not as severe a limitation as for magnetic films. Some progress in improving the sensitivity has recently been observed [43] with Se-SnO heterojunction with applied bias fields. The sensitization is associated with the photocurrent produced by the high electric field in the junction region.

IX. THERMOPLASTICS

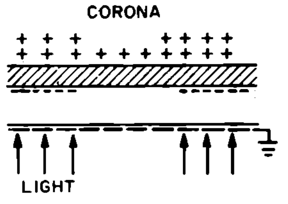
A procedure for recording holograms in thermoplastics is demonstrated schematically in Figure 4 .[44]

① INITIAL CHARGING



② ILLUMINATION

CORONA



③ HEAT TREATMENT

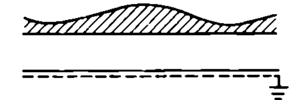


Figure 4 :
Hologram recording with thermoplastic-photoconductor layers [44].

A sandwich layer of a photoconductor (typically doped poly-n-vinyl carbazole) and thermoplastic are deposited on a substrate from solution with a transparent electrode between the substrate and the photoconductor. The thermoplastic is charged with a corona so that a field develops across the sandwich. The potential distribution is modified during illumination when the photoconductor has a low impedance, so that on further charging the electric field distribution across the thermoplastic replicates the optical intensity distribution. Briefly heating the thermoplastic to its softening point results in the relief

pattern shown in Figure 4 due to the contraction of the film in regions of high field. Modulation of the film thickness d provides the phase modulation of the reconstructing beam. The precise control of the charging needed for uniform response has been accomplished [45] by placing a conducting grid at the required potential (about 250 V) close to the thermoplastic , containing apertures for each of the holograms.

Erasure is accomplished by prolonged heating of the film near the softening point when surface tension restores the original form. Conductivity at higher temperatures removes the electric field.

The great advantage of thermoplastics is the high sensitivity for writing, which is comparable to the best photographic emulsions. However, thermoplastics have many drawbacks. The most serious problem is the deterioration of the erasure characteristic. For instance with Staybelite ester 10 erasure is difficult if a hologram has been stored for 24 hours. Even for rapid cycling of the information the signal to noise becomes intolerable after about 500 cycles [45]. The prospect for reducing or eliminating these problems appears to be good because of the large number of thermoplastic materials and properties which are available. Another problem is the limited spatial frequency bandwidth. At high spatial frequencies the field is not uniform in the film, while at low spatial frequencies no concentration occurs. The maximum response is centered close to the spatial frequency $K \sim \pi/d$, so for the high spatial frequencies required for optical holography very thin and uniform films are required (i.e $d < 0.5\ \mu m$) which are difficult to produce experimentally.

X. FERROELECTRIC PHOTOCONDUCTOR

The ferroelectric-photoconductor sandwich shown in Figure 5 has been used for holographic storage, [46] although this system seems to be better suited to the page composer of the optical memory. Recording is accomplished with an electric field across the sandwich in the direction needed to reverse the spontaneous polarization of the ferroelectric.

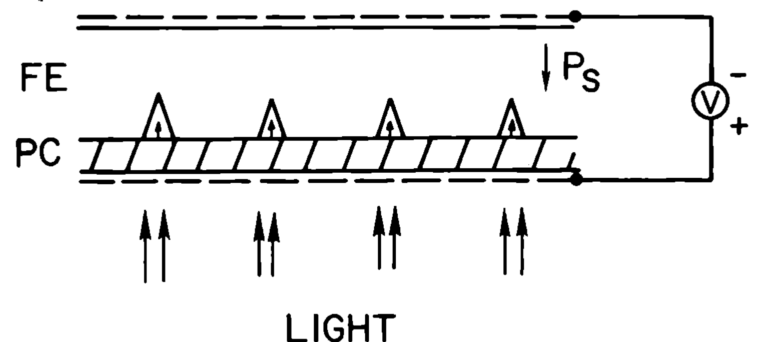


Figure 5 :
Hologram recording with ferroelectric photoconductor layers showing wedge shaped domain formation.

Polarization reversal can only take place under the illuminated regions of the photoconductor which must be sufficiently conducting to allow charge compensation of the depolarizing fields of the reversed domains. In regions where such compensation is not possible (in the dark regions) the polarization cannot reverse. Ferroelectric $Bi_4Ti_3O_{12}$ is well suited for this device since reversal of the polarization along the c axis is accompanied by a rotation of the optical indicatrix in the a-b plane, allowing non-destructive readout due to the birefringence change. Any photoconductor with a high dark resistance and high photoconductive gain, such as ZnSe [46] or CdS/ZnS mixtures [47] is suitable for this application.

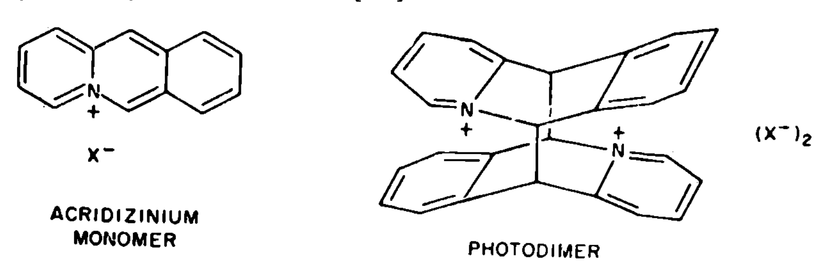
A difficulty of this device is that polarization reversal only takes place near the ferroelectric surface to a depth somewhat greater than the optical fringe spacing of the interference pattern, as shown in Figure 5. This is due to the fringe fields and the nature of the domain growth during polarization reversal. The resolution required for holographic storage means that the domains are extremely small, providing only small retardation of the reconstruction beam. For instance with 500 lines/mm a diffraction efficiency of 10^{-4} was obtained [46]. The recording sensitivity varies rapidly with spatial frequency until the domain length equals the crystal thickness. As with the thermoplastics only very low power densities are required with efficient photoconductors (0.1 to 1 mJ/cm^2).

Since the spatial frequencies required in the page composer are much smaller than the memory plane, this seems a more suitable application.

XI. PHOTOCHROMICS

The term photochromics refers to all materials which undergo a change of optical absorption during irradiation. The change of absorption is necessarily accompanied by a change of refractive index of the medium (fluctuation dissipation theorem) so that either absorption holograms or phase holograms may be recorded in photochromics, depending on whether Δn or $\Delta\alpha$ is dominant at the reconstruction wavelength. The attractive features of photochromics for optical storage is their high resolution, limited in principle by the density of absorbing centers, the lack of any development process and their reversibility (in the materials of interest). A wide variety of photochromic systems have been investigated including ionic crystals such as the alkali halides, photochromic glasses and various organic materials. The microscopic mechanisms vary from one system to another. In ionic crystals and glasses the process involves the optical transfer of electrons between color centers or impurities. In the organics the color change is due to cis-trans isomerism, photoinduced dimerization or polymerization. Numerous examples have been known for many years [48]. However, the net result in all cases is a decrease of the absorption at the recording wavelength and an

increase of the absorption at another wavelength, say λ_e. The materials are returned to their initial state by illumination of the material at wavelength λ_e, thereby erasing the information. As an example of the absorption change the spectra of the monomer and photodimer of acridizinium is shown in Figure 6 together with the molecular structures [49]



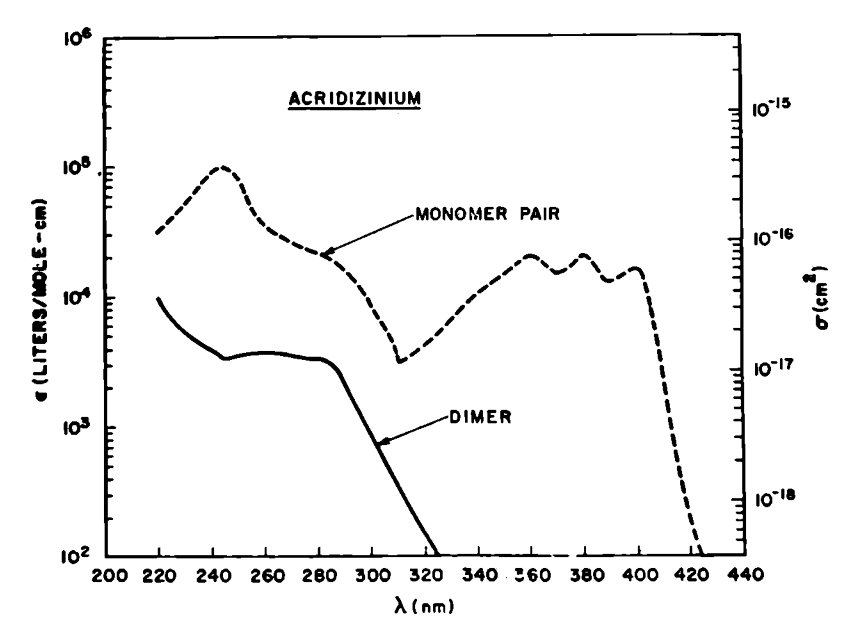


Figure 6 :
Absorption spectra and molecular structure of a monomer and reversible photodimer of acridizinium [49] .

The information is recorded with light of wavelength near 400 nm and erased with light near 300 nm. In most materials the absorption bands of the two states overlap, so that some equilibrium concentration of the two states is reached. Reconstruction of holograms with the same light used for recording is not usually practical since the reconstructing radiation continues to create an absorption change with corresponding degradation

of the recorded information. The use of thick photochromic media for multiple image storage has the additional problem that the optical density changes during exposure, thereby changing the recording sensitivity of the medium for subsequent images. The most practical approach appears to be the use of thin recording media and readout with wavelengths where the material is transparent, making use of the index change which arises either due to the absorption change or to the density change which accompanies photo-induced polymerization [50]. In this way efficient non-destructive readout is possible. For thin photochromic media a high density of photochromic centers is required. For instance a film of 1 μm thickness containing absorbing centers of absorption cross section $\sigma \sim 10^{-17}$ cm^2 requires a concentration of $\sim 10^{21}$ $centers/cm^3$, for efficient absorption at the recording wavelength. Many of the organic crystals investigated satisfy this requirement provided they are not diluted in an inert matrix. [49]

It is possible to estimate an ultimate limit of the sensitivity of photochromic storage for those systems where the absorption and index change are due to localized changes of the electronic configuration. For no overlap of the initial absorption and the optically induced absorption (the optimum case)

$$\Delta\alpha \lesseqgtr \sigma N \qquad (11)$$

where N is the density of absorbed photons, $N \simeq W_0 / h\nu$. An equality in Equation 11 implies a quantum efficiency of unity for the photochromic process. Since $\sigma < 10^{-16} cm^2$ for almost all centers., the minimum incident energy density for a diffraction efficiency of 1% can be estimated from Equations 1 and 11 to be

$$W_0 > 4 \text{ mJ} / \text{cm}^2$$

Alternatively for phase holograms a normalized index change n_0 can be defined as the total index change Δn divided by the concentration of photochromic centers contributing to the change, then :

$$\Delta n < n_0 N \qquad (12)$$

where the equality again implies a quantum efficiency of unity. Now n_0 can be related to the change of molar refraction of the absorbing species ΔR [49] by

$$n_0 = \frac{(n^2+2)^2 \Delta R}{6nA} \qquad (13)$$

where A is Avogadro's number. In order to estimate the maximum value of n_0 off resonance (where σ is small) we can take $\Delta R \sim R \sim 300$ cm^3, then a reasonable upper limit of n_0 is $10^{-21} cm^3$ for small molecules. Using equations 2 and 12 we find for a diffraction efficiency of 1%,

$$W_0 \gtrsim 1 \text{mJ}/\text{cm}^2$$

This estimate may not be valid for the index changes caused by the gross density changes in plastic materials, but is probably an upper limit for photochromic changes in glasses, ionic crystals and the organic systems involving isomerism and dimerism.

Some of the organic systems which have been investigated have quantum efficiencies close to unity [49] and sensitivities about 30 times below the above estimate have been achieved.

It appears that greatly improved sensitivity from photochromics cannot be expected although the sensitivity limits estimated above could be acceptable if combined with long storage lifetimes and no fatigue during repeated recycling. However, with organics fatigue is frequently a problem due to the formation of other products during photoexcitation, and in many photochromics storage lifetimes are limited by the thermal decay of the photoexcited species.

XII. PHOTOREFRACTIVE PROCESS

A system which possesses all the merits of the photochromic process (since it is closely related), yet in principle has a higher sensitivity is the photorefractive process in electro-optic crystals. As with the photochromic process electrons are photoexcited by optical absorption and transferred to a different site. The difference between the two effects is that the change in refractive index resulting from the optical excitation is due to the change of polarization ΔP of the electro optic crystal, as a result of the charge transfer and not directly due to an absorption change or structural change of the crystal. The change of polarization gives rise to the index change via the electro optic effect.

$$\Delta n_i = \frac{n_i^3}{2} r_{ij} \frac{\Delta P_j}{(\varepsilon_i - 1)\varepsilon_0} \qquad (14)$$

r_{ij} is the linear electro optic coefficient of the material and ε_j is the dielectric constant.

This effect has been observed in many electro-optic materials including $LiNbO_3$, $LiTaO_3$ [51], $BaTiO_3$ [52], CdS [53], $Sr_{1-x}Ba_xNb_2O_6$ [54], $BaNaNb_5O_{15}$ [55], $K(Ta,Nb)O_3$ [56], $Bi_4Ti_3O_{12}$ [57] and PLZT ceramics [58]. Since the demonstration of holographic storage in $LiNbO_3$ [59] a great deal of research has been devoted to the understanding and optimizing of this effect for storage applications. Chen proposed that the effect was due to the drift of photoexcited carriers along the polar axis of the crystal to regions of low light intensity, where they were trapped. The resulting space charge field, as shown in Figure 7,

$$E_s = \frac{1}{\varepsilon\varepsilon_0} \int \rho \, dx \qquad (15)$$

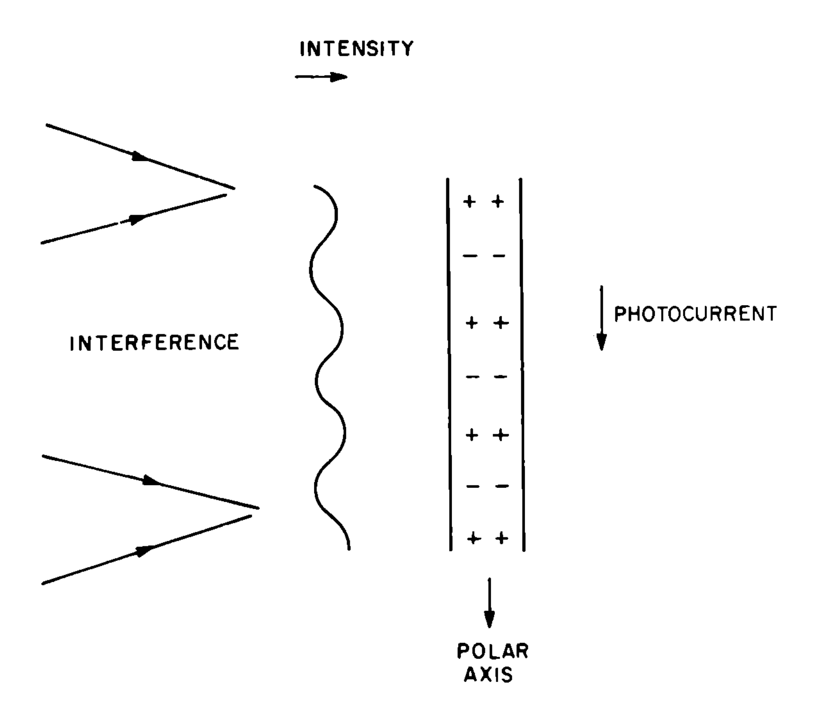


Figure 7 :
Formation of space charged field in photorefractive media due to photo-induced charge separation.

giving the index change according to Equation 14. Erasure is accomplished by uniformly illuminating the crystal with the same wavelength light so that the carriers return to their original sites. In $LiNbO_3$ the process is completely reversible without any indication of fatigue [60]. The primary source of photocarriers for 5145 Å excitation of undoped crystals was found to be small concentration of Fe^{2+} impurities[61] (∿10 ppm) which were presumably converted to Fe^{3+} by intervalence transfer [62] . Deliberately doping the crystals with these impurities was found to greatly enhance the sensitivity, but with a decrease of optical transparency and decreased signal to noise ratio. At the present time under conditions of maximum recording sensitivity (see section II) a diffraction efficiency of 30% is obtained with an incident energy density of approximately 2 J/cm^2 , and this is independent of spatial frequency[63,64]. With some sacrifice of recording sensitivity diffraction efficiencies approaching 100% have been obtained with lower impurity concentration.

Other impurities such as copper, manganese [65] and rhodium [66] give improved sensitivity in $LiNbO_3$ but the figure of merit $S' = \eta_0^{1/2} / \alpha W_0$ appears to be greatest for Fe^{2+} impurities for 5145 Å excitation. The figure of merit has been found to vary with recording wavelength [67] so that the choice of dopant depends on the laser used for the memory system.

Impurities have also proved to play an important role in $BaNaNb_5O_{15}$ [55] and $BaTiO_3$ [68], but in many ferro-electrics the active centers producing the index change have not been identified. Using nominally undoped $Sr_{0.75}Ba_{0.25}Nb_2O_6$ Thaxter [22] has obtained extremely high sensitivity with 4880 Å radiation. He demonstrated that a diffraction efficiency of 1% is produced with an incident energy of about 10 mJ/cm^2 in crystals 0.5 cm thick. With an applied field of 3000 V/cm a further three fold improvement of sensitivity is obtained. Presumable doping crystals to obtain more efficient electronic excitation will give rise to further improvements. However, the storage time of this material is considerably shorter than $LiNbO_3$. Indeed this storage mechanism suffers from the same disadvantages as the photochromic techniques in that thermal decay of the trapped carriers limits the lifetime and reconstruction with the same light used for writing degrades the stored information. However, methods of increasing the stability will be discussed below after a discussion of the microscopic transport processes and the ultimate sensitivity of the photorefractive process.

The photocurrent density J(z) at position z in the crystal due to the optical intensity I(z) is given by

$$J(z) = k\alpha I(z) + \sigma E(z) + eD \frac{dn}{dz} + \Phi \frac{dI}{dt} \qquad (16)$$

where the first term is the bulk photovoltaic effect which is characteristic of pyroelectric crystals [63] due to the asymmetry of the optical charge transfer process. k is determined by the degree of asymmetry, the electronic states involved and the mean free path of the excited carrier. This effect is distinct from other photovoltaic effects involving macroscopic inhomogeneities. The second term is the conduction current, the third term the diffusion current and the last term is the transient current which arises from the polarization change when the light is turned on. This last term, where ϕ is just a proportionality constant includes pyroelectric and excited state dipole effects [69, 74]. Since the build up of the space charge field which gives the index change according to Equation 15 is

$$E_s = \frac{1}{\varepsilon\varepsilon_0} \int J dt \qquad (17)$$

the magnitude of the photocurrent determines the recording sensitivity.

The dominant transport process depends on the material and the manner in which the information is recorded. For instance in

$LiNbO_3$ for cw illumination the photovoltaic current accounts for the index change [63] and saturation occurs when the space field $E_s = k\alpha I/\sigma'$ so that J=0. With external applied fields the conduction current contributes to J [70,71], and increased recording sensitivity is observed. In fact in $K(Ta,Nb)O_3$, PLZT and $Sr_{0.75}Ba_{0.25}Nb_2O_6$ the conduction term is dominant even for quite small fields.[22,56,72]. It appears that diffusion does not play an important role for practical spatial frequencies. Indeed the diffusion current and hence the recording sensitivity would vary with spatial frequency, which is fortunately not the case in practice.

The maximum sensitivity which can be achieved with a recording process of this kind corresponds to the case when the drift length of the photoexcited electrons is comparable to the fringe spacing [73]. Specifically, for a quantum efficiency of unity, the maximum value of the space charge field is given by Equation 15 with

$$\int \rho dx = \frac{\alpha W_0 e\Lambda}{2\pi h\nu} \qquad (18)$$

Now, it has been shown for a wide variety of oxides [74] that within a factor of two, the ratio

$$\frac{r_{33}}{(\varepsilon_3-1)\varepsilon_0} \sim 0.25\ P_3\ m^2/C \qquad (19)$$

where P_3 is the total polarization of the crystal - spontaneous and field induced along the polar axis. Thus using equations 4, 14, 15, 18 and 19 we have all the necessary information to estimate the ultimate sensitivity of the photorefractive process. Taking $n_3^3 \sim 10$ for 0.5 μm radiation we find

$$\frac{\Delta n}{\alpha W_0} < 8\Lambda P_3 \times 10^6\ cm^3/J$$

so that with Equation 1 we find the minimum incident energy density required for a diffraction efficiency of 1% to be

$$W_0 > \frac{2}{\Lambda P_3} \times 10^{-13} J/cm^2$$

where the absorption at the reconstruction wavelength has been neglected. If the recording sensitivity is to be independent of spatial frequency, then the carrier drift length must be less than the smallest fringe spacing of the hologram.

If we take the maximum value of P_3 to be 10^{-4} C/cm² which is greater than but of the same order of magnitude as the spontaneous polarization of many ferroelectric oxides [74], then for a resolution of 1000 lines/mm

$$W_0 > 20\ \mu J/cm^2$$

This sensitivity is comparable to that of holographic emulsions.

It is evident from the analysis that there is not necessarily any advantage in using materials with large electro-optic activity. A large polarization and efficient transport of the photoexcited carriers are the major requirements.

The variations of the space charge field generated by the light will relax with the dielectric relaxation time $\varepsilon_0\varepsilon/\sigma'$ of the material so that for long storage times high resistivity material is required. In high purity $LiNbO_3$ the storage time is of the order of months, but the increased dark conductivity of iron doped crystals leads to decreased storage times. The rapid relaxation reported for instance in $BaTiO_3$ [52] can probably be accounted for by the relatively high conductivity of the material. Uniform illumination of the crystals with the light used for writing creates photocarriers and the increased conductivity results in erasure [71, 63]. However, Johnston [75] has pointed out that in a pyroelectric crystal a polarization variation can occur without a corresponding variation of the electric field provided div.$P_s = \rho$ where the variation of P_s is due to the change in dipole moment of the crystal just due to the change of electronic configuration. So, depending on the transport processes an index change could be present in a crystal following dielectric relaxation due to finite conductivity. In this case the change of dipole moment can only be a few Debye per excited carrier which gives a macroscopic polarization change considerably smaller than that due to macroscopic fields [71], and many more trapped carriers are required.

To avoid erasure during reconstruction of the holograms either a different wavelength of light from that used for recording is necessary, limiting the recording to thin holograms, or multiphoton recording processes can be used (see section XIII), or else the holograms must be made insensitive to erasure in some other way. Increasing the recording sensitivity compared with the erasure sensitivity is possible with applied fields or by suitable doping [70]. This allows many more holograms to be superimposed in the same volume, but does not satisfactorily reduce erasure during reconstruction. Two approaches have been successfully used to fix the holograms by relaxing the electronic space charge fields by ionic motion. One procedure in $LiNbO_3$ involves heating the crystal above 100°C after exposure of the hologram. [76] This results in erasure of the hologram by ionic conductivity but not loss of the information. Illumination of the crystal after cooling to room temperature results in redistribution of the trapped electrons but the ionic pattern remains fixed. Holograms fixed in this way have greatly increased stability, yet they can be erased by heating the crystal above 300°C.

The other procedure involves spatially inhomogeneous reversal of the spontaneous polarization of ferroelectrics in response to the optically generated space charge fields. This may happen

spontaneously (without external fields) in materials of sufficiently low coercivity [22, 72], such as SBN or PLZT. Alternatively, in materials of higher coercivity a uniform external field can be applied in such a manner that the total field in the crystal (external + optically generated) is sufficient to reverse the polarization in the regions of high (or low) light intensity. [72] The resulting polarization is fixed, yet it can be erased by application of an external field above the coercive field of the entire crystal. Several variations of the recording and fixing process have been demonstrated by electric field and temperature control.

The large electric field variation of the recording sensitivity and reconstruction efficiency in SBN [22] and KTN [56] may prove useful for stacked holograms (see section IV). For both these materials the electron transport process is dominated by conductivity, and the effective electro-optic coefficient in Equation 14 is very field dependent because of the large quadratic electro-optic effect near the Curie temperature of these materials.

▪ XIII. MULTI-PHOTON ABSORPTION

A promising approach to avoiding erasure during reconstruction of thick holograms in photorefractive materials by using two step absorption was suggested by Kurz [77]. Consider the energy diagram at the left of Fig. 8 where the ground state and first excited state are bound impurity states while the third state is the conduction band of the host lattice. The holographic information is carried by the radiation corresponding to the bound to bound transition. This alone cannot give charge transfer and hence no permanent storage. Simultaneous irradiation with a uniform beam of energy corresponding to the transition from the excited state to the conduction band can give charge transfer, the efficiency of the excited state absorption depending on the excited state population at each point in the crystal, and the excited state absorption cross section. Thus the recording process is linear and can be read out non-destructively with the same light which carried the information during recording.

Volume holographic storage by two step absorption in copper doped $LiNbO_3$ was performed [78] with a mode locked Nd glass laser. The fundamental wavelength at 1.06 μm corresponds to the bound state transition of the Cu^{2+} ion while the second harmonic (generated with a KDP frequency doubler) corresponded to the excited state transition to the conduction band. Picosecond pulses were required since the lifetime of the Cu^{2+} ion excited state was only 30 psec.[79] A simple interference pattern was created with two 1.06 μm beams at the crystal which was simultaneously illuminated with a single 0.53 μm beam. The information could be read with 1.06 μm radiation alone and erased with uniform illumination with both 0.53 and 1.06 μm radiation. The energy efficiency for recording was approximately five times greater than that for linear absorption in iron doped $LiNbO_3$.

However, the copper was not necessary for storage. Two photon absorption in high purity $LiNbO_3$ with no real intermediate state

absorption had essentially the same efficiency as the copper doped crystal suggesting that even in the doped crystal this process was dominant. Storage was performed using the absorption processes shown on the right of Fig. 8

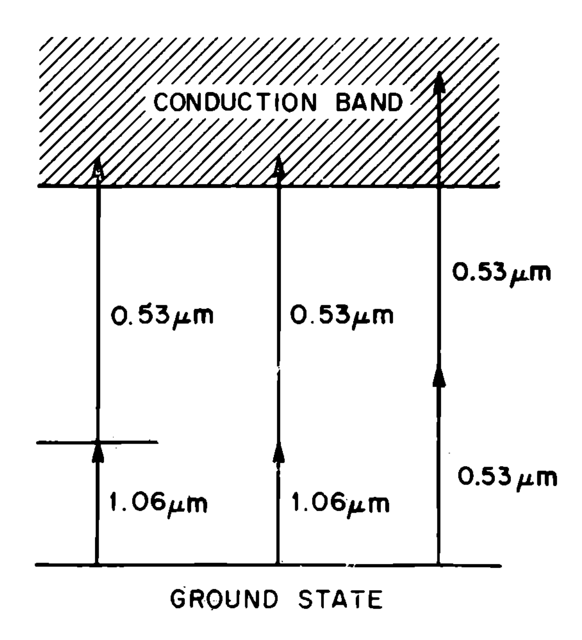


Figure 8 :
Multiphoton absortion processes : left : two-step absorption via a real intermediate state, center : two-photon absorption with a virtual intermediate state with two different photon energies, right : two identical photon energies.

The high purity crystals have the advantage that there is no absorption loss at the reconstruction wavelength and the high resistivity allows long storage times.

Although the efficiency of this process is high, the required intensity of about 100 MW/cm^2 for $LiNbO_3$ is too high to be practical with currently available repetitively pulsed lasers. However these experiments were repeated with undoped KTN crystals with a Curie temperature of 10°C with applied electric fields up to 10^4 V/cm . A large increase of the sensitivity of four orders of

magnitude over $LiNbO_3$ was obtained due to the improved transport processes in this crystal. For instance a diffraction efficiency of a few percent was obtained from a grating of $\sim$10 lines/mm with a single picosecond pulse of 10 psecs duration with an energy of 100 $\mu J/cm^2$ and an intensity of 10 MW/cm^2 . The high peak intensity is necessary since two photon absorption increases quadratically with intensity. However, since the index change integrates with incident energy some trade off of intensity for an increased number of pulses is possible. For instance with a cw mode locked laser of 1W average power and a 4×10^8 repetition rate each of 30 psecs duration, which are readily available give peak intensities of 10 KW/cm^2 over a 1 mm area, so that a diffraction efficiency of 10% can be obtained with KTN in 3 milliseconds (10^6 pulses). The number of pulses and the exposure time decreases quadratically with increased intensity.

A disadvantage of KTN at present is that the storage time of the holograms is only about 10 hours. The conductivity also causes a deterioration of the optical quality of the crystals with an applied field. Hopefully the conductivity can be decreased by improving the crystal purity or by compensation. Alternatively operation at a somewhat lower temperature may be worthwhile. It seems that it may be possible to approach the limiting sensitivity of the photorefractive process with KTN.

Writing by two photon absorption with two different wavelengths, one of energy less than half the bandgap is preferable to two photon absorption with one wavelength as shown on Figure 8, since the recording is linear. With just one wavelength the recording is quadratic in intensity and erasure by two photon absorption during readout is possible, although this is only a serious problem for high intensity read beams.

Finally it should be noted that excitation high into the conduction band, where transport processes may be favorable is possible in thick crystals by two photon absorption. Linear absorption at such wavelengths would only be possible in very thin regions because of the high linear absorption cross section compared to the two photon absorption cross section.

▪ XIV. OTHER APPLICATIONS

The emphasis of this article has been the long term storage of information for high capacity memories. Of course there are a large number of other applications of holography which impose less severe requirements on the recording medium. Elimination of the development step of a holographic emulsion would be very desirable for most applications including holographic interferometry and optical signal processing applications. For such applications long lifetimes may not be required or even desirable. For instance for time average interferometry short integration times, and real time viewing could be useful.

Optical holography could find limited application in integrated optics. Integrated optical circuits will eventually require some kind of memory, if only for switching and logic applications, and it is convenient to have reversible storage in the waveguide itself rather than in the electronic circuitry. It is clear that holographic storage in a two dimensional waveguide could not have the capacity of the three dimensional system. Efficient integrated deflectors which could rapidly address a planar array of holograms are already available.

■ XV. CONCLUSIONS

Over the last few years several systems have been developed which demonstrate the feasibility of holographic recording for mass storage and rapid retrieval of information. At the present time practical systems are limited to read only memories because of the lack of a suitable recording medium for read write applications. There is little point in making an overall comparison of the usefulness of the different materials since each of them satisfies some of the requirements better than others. In fact, one of the most important characteristics of a storage medium, the signal/noise ratio of the reconstructed image, is not available for most of the materials studied. Not until one material satisfies at least the minimum of all of the requirements can progress in the development of a practical read write memory be made. Much of the matrials research over the last few years has been devoted to photorefractive materials probably because the potential of these materials was less clear than the others. These materials have several of the desirable characteristics and the present understanding does raise one's hopes for the development of a practical recording medium, but a great deal of effort is still required to determine the future of this approach.

Helpful comments on this manuscript by L.K. Anderson, D. Von Der Linde and W.J. Tomlinson are gratefully appreciated.

■ REFERENCES

[1] Precision Instruments 'Unicon"

[2] A. Bardos , Appl. Optics 13 , 832 (1974)

[3] H.N. Roberts , J.W. Watkins and R.H. Johnson, Appl. Optics 13 , 841 (1974)

[4] W.C. Stewart, R.S. Mezrich, L.S. Cosentino, E.M. Nagle, F.S. Wendt and R.D. Lohman, R.C.A. Review 34 , 3 (1973)

[5] W.H. Strehlow, R.L. Dennison and J.R. Packard, " The Megafetch Processor " 3M Company, presented at OSA Meeting , April 1974.

[6] L. D'Auria, J.P. Huignard, C. Slezak and E. Spitz, Appl. Opt. 13 , 808 (1974)

[7] H. Kogelnik, Bell System Tech. J. 48 , 2909 (1969).

[8] L.H. Lin, J. Opt. Soc. Am. 61,203 (1971)

[9] G.R. Knight, Appl. Opt. 13 , 904 (1974).

[10] K. Tsukamoto, A. Ishii, A. Ishida, M. Sumi and N. Uchida Appl. Opt. 13 , 869 (1974).

[11] O.N. Tufte and D. Chen, I.E.E.E. Spectrum, Feb. 1973 p. 26

[12] A.W. Smith, Appl. Opt. 13, 795 (1974).

[13] B.R. Brown , Appl. Opt. 13 , 761 (1974)

[14] J.D. Zook, Appl. Opt. 13, 875 (1974)

[15] G. Labrunie , J. Robert and J. Borel , Appl. Optics 13, 1355 (1974).

[16] J.R. Maldonado and A.H. Meitzler, I.E.E.E. Trans. ED-17 148 (1970)

[17] G.W. Taylor , Ferroelectrics 1, 79 (1970).

[18] Y. Takeda , Appl. Opt. 13, 825 (1974).

[19] W.C. Stewart and L.S. Cosentino , Appl. Opt. 9, 2271 (1970,

[20] J.A. Rajchman , Appl. Opt. 9, 2269 (1970).

[21] A. Van Der Lugt, Appl. Optics 12 , 1675 (1973)
L.K. Anderson, I.E.E.E. Trans. MAG-7 , 601 (1971)

[22] J.B. Thaxter and M. Kestigian , Appl. Opt. 13 , 913 (1974)

[23] D. Pohl, Appl. Optics 13, 341 (1974)

[24] L. D'Auria , J.P. Huignard and E. Spitz, I.E.E.E. Trans. MAG 9, 83 (1973).

[25] L.H. Lin, Appl. Opt. 8, 963 (1969)

[26] R.A. Bartolini, Appl. Opt. 13, 129 (1974)

[27] R.G. Zech, J.C. Dwyer, H. Fichter and M. Lewis, Appl. Opt. 12 , 2822 (1973).

[28] D. Chen , J.F. Ready and E. Bernal , J. Appl. Phys. 39, 3916 (1968).

[29] R.S. Mezrich, Appl. Optics 9, 2275 (1970).

[30] G.Y. Fan and J.G. Greiner , J. Appl. Phys. 39 , 1216 (1968) ; 41 , 1401 (1970).

[31] J.T. Chang , J.F. Dillon and U.F. Gianola , J. Appl. Phys. 36 , 1110 (1965).

[32] P. Coeuré, J.C. Gray and J. Carcay, I.E.E.E. Trans. MAG-7 397 (1971).

[33] R.E. MacDonald and J.W. Beck , J. Appl. Phys. 40, 1429 (1969)

[34] P. Chaudhari , J.J. Cuombo and R.J. Gambino , Appl. Phys. Lett. 22 , 337 (1973).

[35] D. Chen , Appl. Optics 13 , 767 (1974).

[36] R.G. Brandes, F.P. Laming , A.D. Pearson, Appl. Optics 9, 1712 (1970).

[37] Y. Ohmachi and T. Igo, Appl. Phys. Lett. 20 , 506 (1972).

[38] S.A. Keneman , Appl. Phys. Lett. 19 , 205 (1971).

[39] J. Feinleib, J. de Neufville , S.C. Moss and S.R. Ovshinsky, Appl. Phys. Lett. 18 , 254 (1970).

[40] R.J. Von Gutfeld and P. Chaudhari , J. Appl. Phys. 43 , 4688 (1972).

[41] A.D. Pearson and B.G. Bagley, Mat. Res. Bull 6, 1041 (1971).

[42] A.F. Wells, "Structural Inorganic Chemistry", Oxford Univ. Press, 1962 , p. 682.

[43] M. Okuda , T. Matsushita, T. Yamagami and K. Yamamoto, Appl. Opt. 13, 799 (1974).

[44] "Optical Holography" by R.J. Collier, C.B. Burkhardt and L.H. Lin, Academic Press, N.Y. (1971).

[45] T.C. Lee, Appl. Optics. 13 , 888 (1974).

[46] S.A. Keneman, A. Miller, and G.W. Taylor, Ferroelectrics 3, 131 (1973).

[47] D.B. Fraser, Proc. I.E.E.E. , July 1973.

[48] G.G. Dorion and A.F. Wiebe, "Photochromism" , The Focal Press, London 1970.

[49] W.J. Tomlinson, E.A. Chandross , R.L. Fork, C.A. Pryde and A.A. Lamola , Appl. Optics 11 , 533 (1972).

[50] M.J. Bowden, E.A. Chandross and L.P. Kaminow, Appl. Optics 13, 112 (1974).

[51] A. Ashkin, G.D. Boyd, J.M. Dziedzic , R.G. Smith, A.A. Ballman and K. Nassau, Appl. Phys. Lett. 9, 72 (1966)

[52] R.L. Townsend and J.T. LaMacchia, J. Appl. Phys. 41 5188 (1970).

[53] A. Ashkin, B. Tell and J. Dziedzic, I.E.E.E. J. Quant. El. QE3 , 400 (1967).

[54] J.B. Thaxter , Appl. Phys. Lett. 15 , 210 (1969).

[55] J.J. Amodei, D.L. Staebler and A.W. Stephens, Appl. Phys. Lett. 18 , 507 (1971).

[56] F.S. Chen , J. Appl. Phys. 38 , 3148 (1967).

[57] L.H. Lin, Proc. I.E.E.E. 57 , 252 (1969).

[58] F. Micheron , A. Hermosin, G.B. Smith and J. Nicolas, C.R. Acad. Sci. 8, Dec. 1971.

[59] F.S. Chen , J. Appl. Phys. 40, 3389 (1969).

[60] T.K. Gaylord, T.A. Rabson, F.K. Tittel , Appl. Phys. Lett. 20 , 47 (1972).

[61] G.E. Peterson , A.M. Glass and T.J. Negran, Appl. Phys. Lett. 19 , 130 (1971), A.M. Glass, G.E. Peterson and T.J. Negran , Proc. Symp. on Damage in Laser Materials, NBS Spec. Pub. 372 (1972).

[62] M.G. Clarke, F.J. Di Salvo, A.M. Glass and G.E. Peterson, J. Chem. Phys. 59 , 6209 (1973).

[63] A.M. Glass , D. Von Der Linde and T.J. Negran , Appl. Phys. Lett., Aug 15 (1974).

[64] H. Kurz (private communication).

[65] W. Phillips , J.J. Amodei and D.L. Staebler, R.C.A. Rev. 33 , 94 (1972).

[66] A. Ishida , O. Mikami, S. Miyazawa and M. Sumi, Appl. Phys. Lett. 21, 193 (1972).

[67] H. Kurz, 3rd Int. Meeting on ferroelectricity, Edinburgh 1973 , to be published in "Ferroelectrics".

[68] F. Micheron and G. Bismuth, J Phys. Suppl. to n°4 33 , C2-147 (1972).

[69] A.M. Glass and D.H. Auston , Opt. Comm. 5, 45 (1972)

[70] D.L. Staebler and W. Phillips , Appl. Opt. 13, 788 (1974).

[71] J.J. Amodei and D.L. Staebler , R.C.A. Rev. 33 71 (1972).

[72] F. Micheron , C. Mayeux and J.C. Trotier, Appl. Opt. 13 , 784 (1974).

[73] L. Young, W.K.Y. Wong, M.L.W. Thewalt and W.D. Cornish, Appl. Phys. Lett. 24 , 260 (1974).

[74] S.H. Wemple, M. DiDomenico and I. Camlibel, Appl. Phys. Lett. 12 , 209 (1968).

[75] W.D. Johnston, Jr., J. Appl. Phys. 41 , 3279 (1970).

[76] D.L. Staebler and J.J. Amodei , Ferroelectrics 3 , 107 (1972).

[77] Private communication.

[78] D. Von Der Linde, A.M. Glass and K.F. Rodgers, Appl. Phys. Lett., Aug 1, 1974.

[79] D. Von Der Linde , D.H. Auston and A.M. Glass, Solid State Comm. 14, 137 (1974).

WAVELENGTH DEPENDENCE OF THE PHOTOREFRACTIVE PROCESS IN DOPED $LiNbO_3$

H. Kurz

Philips Forschungs Laboratorium Hamburg GmbH

Hamburg 54, Germany

I. INTRODUCTION

The investigations of the physical properties of doped $LiNbO_3$ are stimulated by promising application in the field of optical storage. The photorefractive process, e.g. the changes in the indices of refraction during illumination with laser light [1] can be used as a basis of an erasable pure phase holographic storage medium [2] and may find future application in other real time holographic devices, distributed feed back systems and integrated optics. Investigations of the spectral dependence of the photorefractive process in doped $LiNbO_3$ appear important because, by comparing the spectral sensitivity of this process and the normal absorption spectra a clear distinction between inefficient and efficient transitions should be possible. This information could be useful to clarify the basic exitation phenomena in photorefractive crystals. The identification of particular transitions responsible for the photorefractive effect is based on recent theoretical and experimental work concerning the absorption spectra of transition metals in $LiNbO_3$ [3,4].

II. EXPERIMENTAL

Single crystals of doped $LiNbO_3$ were grown by Dr. A. Raüber, IAF Freiburg, using the Czochralski method. After pulling a nominal pure crystal the oxides of different transition metals were added to the melt. The doping concentration was monitored by X-ray fluorescence.

The impurity concentration of the crystals other than the major dopants, measured by a semiquantitative spectrochemical analysis, was (wt %) :

Mg < .003	Si < .001	Ni < .002
Pt < .001	Mn < .0001	Co < .003
Cr < .001	Ti < .001	

The measurements of the induced index change, Δn_e, have been performed by conventional holographic technique [2]. The sensitivity of the photorefractive process, $\Delta n_e^{633}/\Delta E$, was measured by He-Ne-laser light at λ = 633 nm as the slope in the recording curve Δn_e^{633} (E) in the initial stage of illumination. The exposure energy density E inducing the index change was provided by argon laser lines between 458 nm and 528 nm.

III. EXPERIMENTAL RESULTS

Various dopants have been investigated. Of those chromium and cobalt seem to be inefficient in the observed spectral region. The apparent absence of a true photorefractive effect of these ions must be either due to the absence of a suitable excitation process at $\lambda > 458$ nm or neutralizing the activation by other residual impurities.

Crystals doped with Ni show highly nonlinear recording behaviour -e.g. slope and saturation value depend on the incident light power- which will be the subject of a special investigation.

For applications Fe, Cu and Mn doping appear particularly attractive, because crystals doped with these impurities exhibit increased sensitivity and high index changes. [4]

Therefore, the results presented here will be restricted to Fe, Cu and Mn doped crystals, whose photosensitivity spectra are shown in Fig. 1 . For comparison of the effects of various dopants the data are fitted in relative units to a common point at λ =458nm. The shape and position of the photorefractive band depend specifically on the kind of doping . This experimental result confirms conclusive the active role of the 3d-ions as sources of photoelectrons [4, 5,6]. The sharp increase of the photorefractive band in $LiNbO_3$: Mn is masked by other inefficient transitions [3]. This may be an explanation of the controversy about the activation by Mn- centres [4,5] . However, in $LiNbO_3$: Fe and $LiNbO_3$: Cu the shape of the photorefractive bands follow the broad absorption in the visible.

The cut-off of the photorefractive sensitivity at the long energy side can be observed at λ = 520 nm for $LiNbO_3$: Mn in this experiment already. From preliminary investigations performed with He-Ne-laser illumination we deduce that in the case of Cu-doping the cut-off lies at $\lambda < 633$ nm. However, the remarkable sensitivity of $LiNbO_3$: Fe in the red and the soft decrease in the spectral sensitivity at longer wavelengths leads to the assumption, that the cut-off of the broad photorefractive band introduced by iron-doping is located at $\lambda > 633$ nm.

The absolute values of the sensitivity and the quantum efficiency -e.g the index change per absorbed photon - are shown in Table 1. Caused by the enhanced optical density at λ = 458 nm

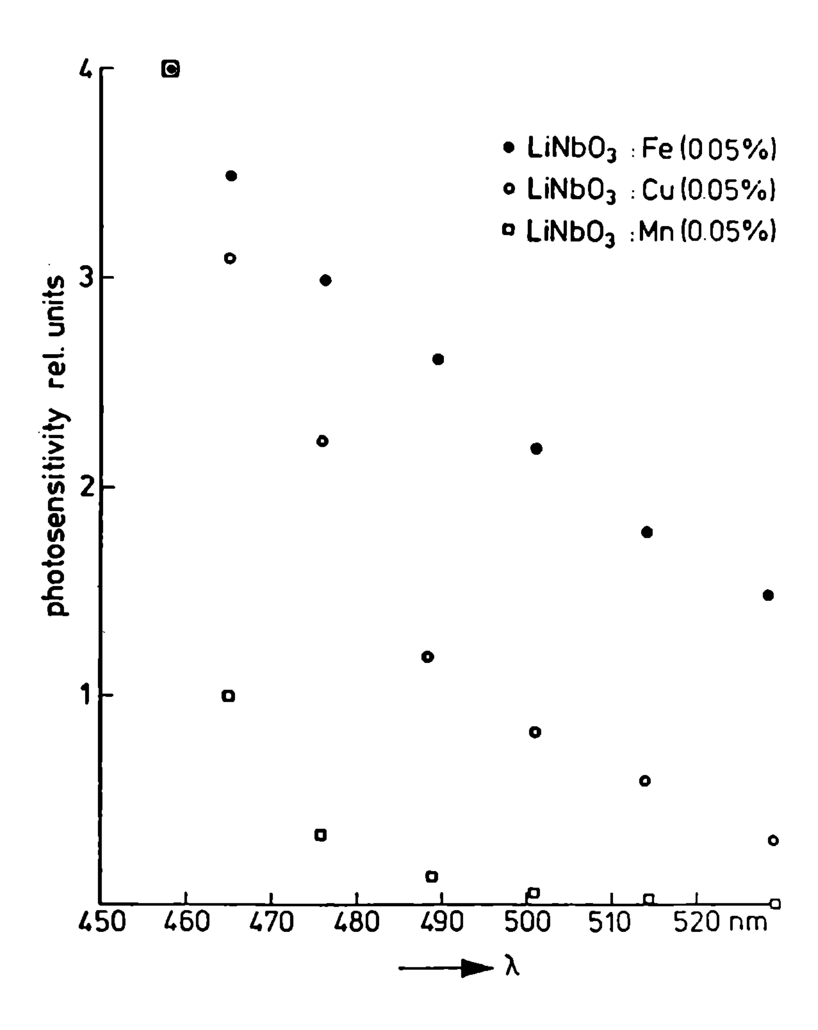


Figure 1 :
Photorefractive sensitivity, $\Delta n^{633}/\Delta E$, vs wavelength, $LiNbO_3$ doped with 0.05 wt% Fe, Cu and Mn.

TABLE 1

Absolute values of the photorefractive sensitivity and the quantum efficiency in doped $LiNbO_3$ at λ=458 nm.

Sample	Sensitivity	Quantum efficiency
$LiNbO_3$: Fe (0.05%)	2.10^{-5} cm^2/J	16.10^{-25} cm^3/photon
$LiNbO_3$: Cu (0.05%)	4.10^{-5} cm^2/J	6.10^{-25} cm^3/photon
$LiNbO_3$: Mn (0.05%)	4.10^{-6} cm^2/J	3.10^{-26} cm^3/photon ★

★) The photorefractive band is covered by other inefficient transitions.

the Cu-doped samples become more sensitive than the Fe-doped crystals. However, the quantum efficiency of $LiNbO_3$:Fe is superior in the whole spectral region observed in this experiment.

We conclude from the different shapes of the spectral dependence that iron contamination alone cannot explain the high index changes in $LiNbO_3$:Cu and $LiNbO_3$:Mn [5]. The contribution of residual Fe-impurities in these crystals to the photorefractive process at longer wavelengths, deduced from the He-Ne-laser experiments mentioned above, was less than 5%.

Additional information about the nature of the photorefractive center can be obtained by measuring the sensitivity gap between reduced and oxidized crystals. [4]

Reduction of the crystal was accomplished by annealing in argon atmosphere for one hour to 650 °C, oxidation by annealing in oxygen atmosphere for five hours to 800°C. Inspection of the sensitivity gaps shown in Fig. 2 and Fig. 3 indicate changes in the density of sensitizing complexes, induced by reducing annealing procedures.

■ IV. DISCUSSION

The sensitivity ratio of the reduced and oxidized samples depends on the wavelength of the laser lines in these experiments.

In previous models the photoinduced charge transfer from Fe^{++} ions to the conduction band was proposed as basic mechanism of the photorefractive process [4,6]. The change in sensitivity then can be related to the variation of the Fe^{++} -ion concentration, induced by chemical treatments. This model requires a constant sensitivity ratio in the observed spectral region.

The measured wavelength dependence of this ratio is difficult to reconcile with this model. Further experiments on extended spectral range should clarify this point.

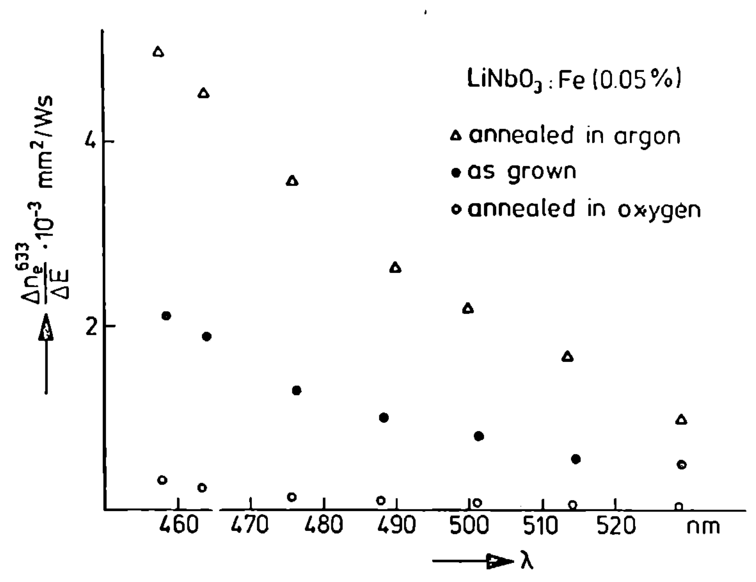


Figure 2 :
Photorefractive sensitivity vs. wavelength,
♦♦♦ *$LiNbO_3$:Fe annealed in argon for I^h/650°C*
●●● *$LiNbO_3$:Fe annealed in oxygen for 5^h/800°C*
ooo *$LiNbO_3$:Fe as grown. All samples are doped with 0,05 wt% iron.*

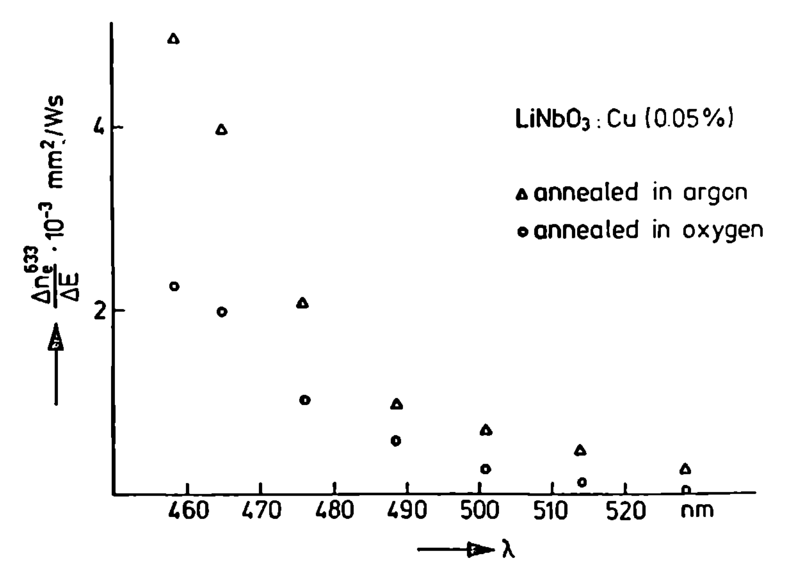


Figure 3 :
Photorefractive sensitivity vs. wavelength,
♦♦♦ *$LiNbO_3$:Cu annealed in argon for I^h/650 °C*
ooo *$LiNbO_3$:Cu annealed in oxygen for 5^h/800 °C*
The samples are doped with 0.05 wt% copper.

▪ REFERENCES

[1] A. Ashkin et al., Appl. Phys. Letters. 9 (1966) p. 72

[2] F.S. Chen, J.T. Macchia and D.B. Fraser, Appl. Phys. Letters 13 (1968), p. 223.

[3] B. Dischler, J.R. Herrington, A. Räuber and J. Schneider, Bull. Am. Phys. Soc. 18 (1973) and to be published.

[4] W. Philips, J.J. Amodei and D.L. Staebler, R.C.A.-Review 33 (1972)

[5] A.M. Glass, G.E. Peterson and T.J. Negran, ASTM-NBS Symposium on Damage in Laser Materials, Boulder, Colorado, June 1972.

[6] G.E. Peterson, A.M. Glass and T.J. Negran, Appl. Phys. Letters 19 (1971), p. 130

SYSTEMES DE STOCKAGE OPTIQUE

E. SPITZ
Thomson CSF
Domaine de Corbeville
91401 Orsay

On se propose de présenter une description simplifiée de systèmes optoélectroniques dans lesquels est réalisée la fonction mémoire. Dans le premier système, on utilise le stockage de l'information sous forme holographique, tandis que dans le second système, l'enregistrement est fait point par point à la surface d'un disque de matière plastique transparente. Dans le cas des mémoires pour ordinateur, on ne discutera pas ici le problème de la comparaison du rapport performance / prix des systèmes de mémoires optiques avec celui d'autres mémoires plus couramment utilisées. Signalons juste que même si les recherches actuelles sur les matériaux d'enregistrement optique laissent prévoir des améliorations importantes de leurs performances, les autres procédés feront eux aussi des progrès dans les années à venir. Il est difficile de prévoir la place future de tels systèmes à moins qu'il y ait des changements importants dans la façon de traiter l'information.

▪ I. MEMOIRE HOLOGRAPHIQUE

L'intérêt des mémoires optiques, qu'elles soient holographiques ou non, réside principalement d'une part dans la très grande densité d'informations que l'on peut en principe inscrire dans un matériau photosensible, puisqu'il suffit en principe d'une région dont les dimensions, limitées par la diffraction, sont de l'ordre de la longueur d'onde de la lumière utilisée, d'autre part, dans la grande vitesse d'accès à l'information puisque l'on sait dévier ou moduler les faisceaux optiques à

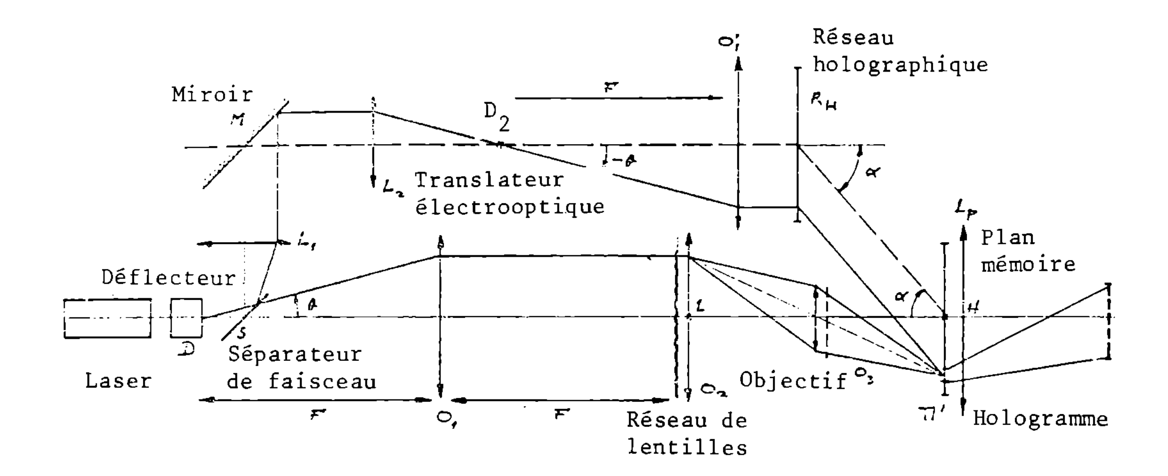


Figure 1:
Schéma de principe d'une mémoire holographique à trois dimensions. La sélection des différents hologrammes contenus dans le milieu d'enregistrement se fait à l'aide du déflecteur D et du translateur électrooptique D_2.

grande vitesse. Des études théoriques ont montré que pour obtenir une mémoire optique à accès aléatoire, inscriptible et effaçable et dont la capacité puisse atteindre 10^{11} à 10^{12} bits avec un débit de lecture de l'ordre de 0,1 à 1 Gbit/s, il était avantageux de considérer l'organisation suivante: on divise l'information en pages de 10^3 à 10^4 bits que l'on enregistre sous forme holographique dans le matériau d'enregistrement. Ce matériau devra être épais et devra permettre la superposition en volume de différents hologrammes que l'on pourra distinguer à la lecture en modifiant la direction du faisceau de détection.

On donne sur la figure 1 le schéma de principe d'une telle mémoire holographique tridimensionnelle, dont une maquette a été réalisée au laboratoire.

Le principe de l'enregistrement est le suivant. On fait interférer à l'intérieur du "plan mémoire" deux faisceaux issus d'un même laser. L'un d'eux, le faisceau de référence passe par les lentilles L_1, L_2, O'_1. L'autre, le faisceau de données, passe par O_1, O_2, O_3. Immédiatement après la lentille O_3 se trouve le "composeur de page" qui permet d'inscrire sur le faisceau optique N bits d'information disponibles à l'entrée de la mémoire. La sélection d'un hologramme particulier à l'intérieur du milieu d'enregistrement se fait de deux façons:
i) en changeant le point d'impact des faisceaux sur le plan mémoire, ce que l'on réalise à l'aide du déflecteur D situé

juste à la sortie du laser. On dispose d'un nombre de zones utiles égal au nombre M de petites lentilles contenues dans le réseau de lentilles situé près de O_2.
ii) en faisant varier l'angle d'incidence sur le plan mémoire du faisceau de référence à l'aide du translateur électrooptique D_2. On obtient ainsi la superposition de P hologrammes dans la même zone de la mémoire. Cela permet d'enregistrer N x M x P bits d'information dans le volume d'enregistrement. Inversement à la lecture, où l'on ne se sert que du faisceau de référence, on choisit l'un des M x P hologrammes enregistrés à l'aide du déflecteur D et du translateur D_2. La lentille L_p sert alors à reconstituer une page d'information que l'on pourra décoder à l'aide d'un réseau de détecteurs, ou utiliser telle quelle pour un traitement optique de l'image.

La maquette d'un tel système a été réalisée en utilisant certains composants décrits ci-dessous. Le déflecteur D est un déflecteur acoustooptique bidirectionnel de trois microsecondes de temps d'accès avec un rendement global de déflexion de l'ordre de 40%. La résolution spatiale de 220 points est compatible avec les (20 x 20) lentilles de Fresnel obtenues par pressage. Le translateur D_2 était constitué dans une première étape par une épaisse lame de verre tournante, mais pourrait être remplacée par un autre déflecteur acoustooptique de temps d'accès beaucoup plus court.

Le composeur de page utilisé est une matrice de cristaux liquides fonctionnant par biréfringence. Bien que d'autres solutions soient en cours d'étude pour la réalisation d'imageurs performants, il a semblé que, indépendamment des temps d'accès, les cristaux liquides constituent actuellement une solution au problème de l'introduction de données dans une maquette de mémoire holographique inscriptible et effaçable. Le mode de fonctionnement utilisé est l'effet de biréfringence controlé dans les cristaux liquides nématiques, dont le temps d'accès est de quelques dizaines de millisecondes.

Le plan mémoire est constitué par une série de cristaux de $LiNbO_3$ dopés avec 0,05% de fer. Chaque cristal est relié à des électrodes permettant l'effacement par élévation locale de la température.

Le plan de lecture est constitué par une matrice de 16 x 16 photodiodes au silicium associées à des transistors MOS de commutation.

Bien que les matériaux photoferroélectriques disponibles actuellement ne permettent pas de réaliser autre chose que des maquettes de démonstration, la construction d'un tel système a permis de résoudre la plupart des problèmes annexes posés par la réalisation d'une mémoire holographique. Il est intéressant de noter que l'on peut réaliser à partir d'une telle maquette un système de mémoire morte très intéressante. Si l'on utilise le système seulement à la lecture et qu'on a préalablement enregistré un très grand nombre d'hologrammes dans un matériau photosensible adéquat (film par exemple), on pourra

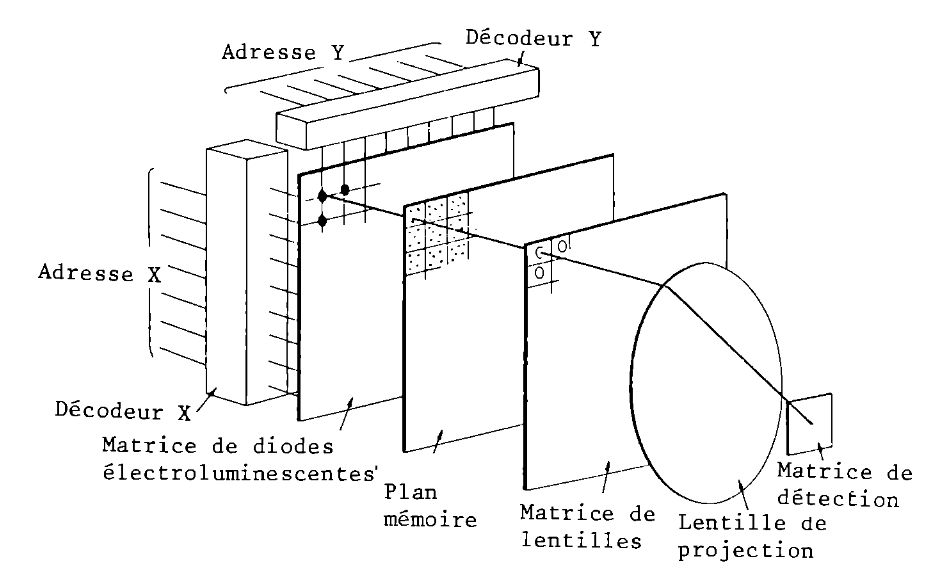


Figure 2:
Schéma de principe d'une mémoire optique morte à accès aléatoire réalisé à l'aide d'un réseau de diodes électroluminescentes.

obtenir une mémoire de très grande densité d'information, par exemple 100 fois plus grande qu'avec les bandes magnétiques actuelles. Signalons que l'on peut envisager un certain nombre d'autres solutions pour l'organisation des mémoires optiques. Parmi celles-ci, nous allons décrire un système à accès rapide où l'adressage se fait par une matrice de diodes électroluminescentes.

Au lieu d'avoir un seul laser suivi par un déflecteur à accès aléatoire, on peut utiliser un réseau de diodes disposées dans un plan et que l'on peut sélectionner par un circuit d'adressage X-Y convenable. La lumière émise par une diode passe alors à travers une microimage où l'information se trouve enregistrée sur film photographique, comme on peut le voir sur la figure 2. On fait alors l'image sur un réseau de diodes de la microimage à l'aide d'une lentille comprise dans un réseau de lentilles. Un tel dispositif a une capacité relativement limitée (de 10^4 à 10^5 bits) mais elle présente certains avantages en ce qui concerne le temps d'accès et la facilité de reproduction ainsi que le faible coût des enregistrements.

■ II. VIDEODISQUE

On va décrire maintenant un système de mémoire optique morte permettant d'enregistrer sur la taille d'un disque audio habituel le signal vidéo correspondant à un programme de télévision en couleur durant une demi-heure. En raison de la très grande quantité d'information à restituer, il est hors de question d'utiliser un système mécanique d'aiguille suivant un sillon comme pour les disques audio. On peut évidemment faire appel à d'autres effets physiques de type mécanique ou électrique, mais il apparait que l'utilisation d'un senseur optique est préférable à bien des égards. Par ailleurs, ce système constitue l'une des premières grandes applications de la photonique, d'où l'exposé détaillé qui va suivre.

Comme il serait difficile d'enregistrer le signal vidéo sous forme analogique, on préfère le traiter en modulation de fréquence, ce qui permet alors de se contenter d'enregistrer les instants où le signal MF passe par zéro. Cela nécessite d'enregistrer une plus grande quantité d'information, soit $1,5 \times 10^{10}$ bits pour un programme de 20 minutes, mais cela permet une plus grande souplesse à la détection puisqu'il s'agit d'un système d'enregistrement à deux états. En prenant un disque tournant à 1500 tours par minute, avec une distance entre spires de 2,5 microns, on obtient environ 23 minutes d'enregistrement. Noter que la vitesse de rotation a été choisie pour qu'un tour donne une image, ce qui permet certaines possibilités dont on parlera plus loin. Avec une telle vitesse de rotation, la résolution spatiale des informations inscrites le long d'un spire compatible avec une bande passante vidéo

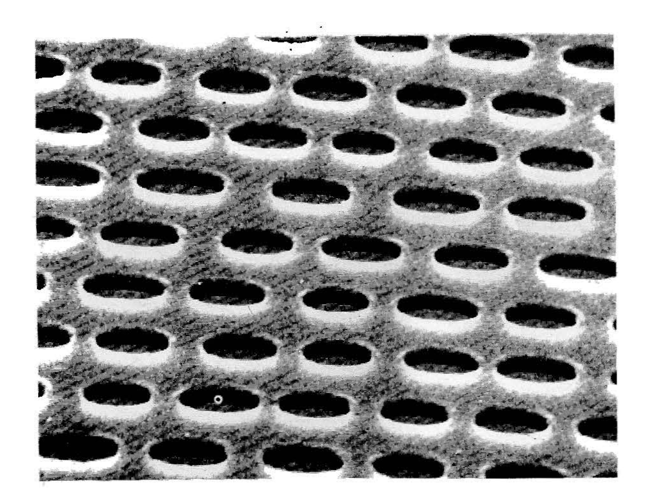


Figure 3:
Exemple de portion d'un enregistrement sur vidéodisque examinée à l'aide d'un microscope électronique à balayage.

de 3 MHz est donc de 0,8 micron environ pour la spire de diamètre le plus petit. On représente sur la figure 3 un exemple de portion d'un disque où l'on aperçoit une série de "cuvettes".

On voit à partir de ces indications numériques sur la finesse des informations qu'il va falloir utiliser d'une part un faisceau fortement focalisé, ce qui va conduire à une très faible profondeur de champ, d'autre part un système de poursuite de la spire permettant au faisceau d'être guidé latéralement à beaucoup mieux que 2,5 microns. Cela a conduit à mettre au point des systèmes d'asservissement de qualité suffisante pour répondre aux exigences précédentes.

La source optique utilisée est un laser He-Ne de un mW focalisé à l'aide d'un objectif de microscope d'ouverture numérique 0,45 donnant une tache focale de diamètre 1,4 microns, avec une profondeur de champ acceptable de 2 microns. Le positionnement vertical du disque qui est fait en matière plastique souple et transparente (PVC ou mylar) de 0,15 mm d'épaisseur est obtenu par effet aérodynamique de l'écoulement de l'air autour de pièces métalliques de formes convenables (stabilisateur en forme de U et dièdre de tranquillisation; figure 4). L'asservissement radial du pinceau lumineux le long de la spirale enregistrée est réalisée par déflexion, au moyen d'un miroir vibrant situé immédiatement en amont de l'objectif. Figure 5.

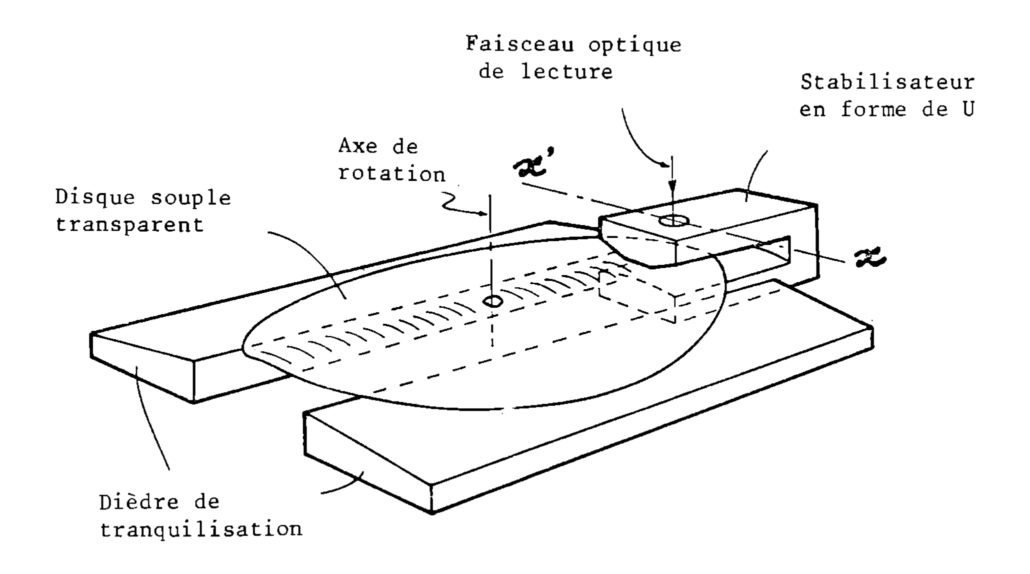


Figure 4:
Dispositif mécanique permettant de maintenir le disque à une distance constante de l'objectif de microscope servant à focaliser le laser.

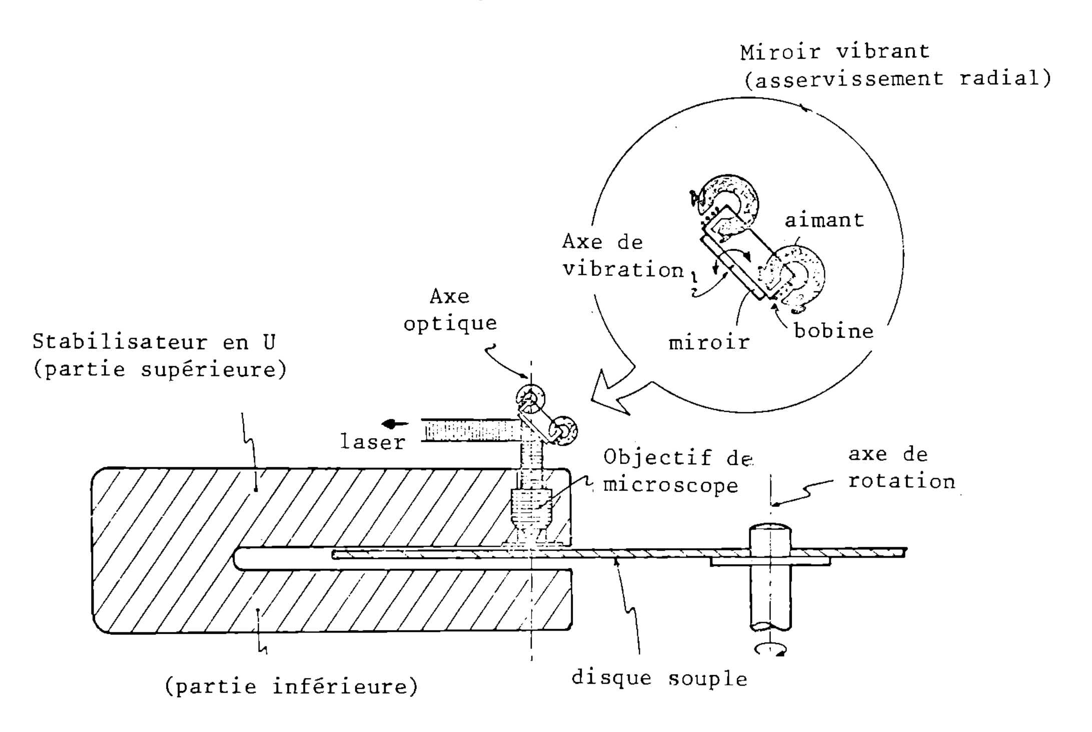


Figure 5:
Dispositif d'asservissement permettant au faisceau laser de suivre la spirale enregistrée.

On utilise le signal HF lui-même pour élaborer les signaux d'erreur nécessaires pour commander la position du miroir. Pour cela on a disposé deux cellules photoélectriques sous le disque symétriquement par rapport au plan constitué par le faisceau laser et la direction moyenne de la spire. On trouve alors qu'il y a dissymétrie entre les signaux fournis par les deux cellules lorsque l'axe du faisceau n'est pas situé sur l'axe de la spire, d'où un signal d'erreur utilisable pour l'asservissement.

On utilise un système de deux autres diodes pour réaliser la lecture, comme il est expliqué sur la figure 6. Le codage de l'information s'apparente au procédé SECAM.

Comme on a choisi la vitesse de rotation de façon à avoir un tour par image, il est facile de réaliser des effets spéciaux tels que arrêt sur une image, ralentis, marche arrière. En ce qui concerne la fabrication des disques, on prépare un premier disque en résine photosensible éclairé par un laser convenablement modulé. On développe ensuite la résine, puis on fabrique une matrice métallique qui servira à faire les disques utilisés

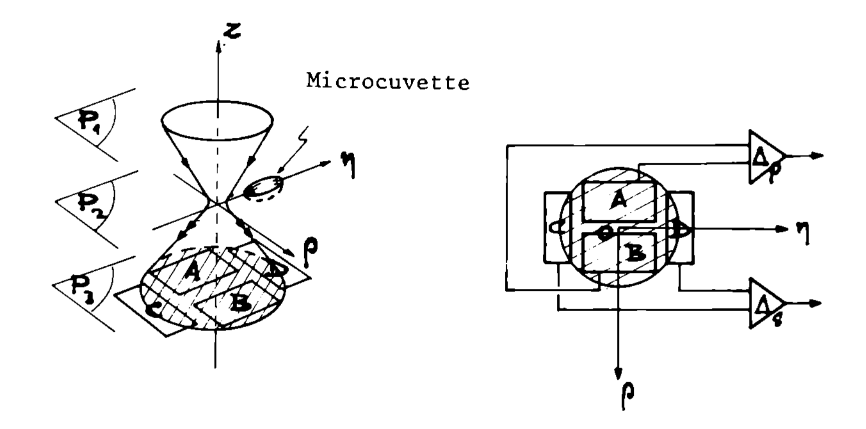


Figure 6:
Schéma indiquant la disposition des cellules photoélectriques utilisées pour détecter le positionnement latéral du faisceau laser par rapport à la spirale, et pour obtenir le signal vidéo. La direction de déplacement des cuvettes est donnée par η *. La différence entre les signaux provenant de A et B sert de signal d'erreur, la différence entre les signaux provenant de C et D donne le signal vidéo.*

par pressage, procédé qui permet de reproduire avec une très bonne fidélité les reliefs jusqu'à quelques centaines d'angströms.

On a discuté dans ce chapitre différents systèmes de mémoires optiques développées récemment dans notre laboratoire. Les performances obtenues permettent d'affirmer que les systèmes optiques vont jouer un rôle important dans le domaine des mémoires mortes grâce à la très forte densité d'information et aussi à la facilité de reproduction soit des films photographiques, soit des surfaces de lames planes obtenues par pressage. En revanche le développement de mémoires inscriptibles et effaçables apparait plus lointain, parce qu'il dépend essentiellement de la possibilité d'obtenir des matériaux ayant des performances très supérieures à celles des matériaux dont on dispose actuellement.

SOLID STATE OPTICAL PROCESSING DEVICES

F. Micheron

Thomson - C.S.F.

91401 ORSAY France

I. INTRODUCTION

Optical processing is a general technique which allows analogical or digital computing with images, i.e. mathematical operations on data contained in one or several images. Analogical operations are for instance sums, differences, spatial differentiation, spectral analysis and correlation ; digital operation are inversion, A N D, O R and Exclusive OR. When the optical device allows two of these basic operations, any Boolean operation can be performed by combining such devices.

A photographic transparency may be considered as a 10^5 - 10^7 bits memory plane, in which information is available as a whole and transmissible in parallel to an optical processor. For instance, the intrinsic processing time, in an one meter long correlator is equal to the light transit time $\tau \simeq$ 30 ns ; accordingly, the intrinsic processing rate for one image is typically 10^{13}- 10^{15} bits/second. This high speed processing constitutes the fundamental advantages of the optical techniques. Nevertheless, processing rate of the system, including data input and output transducersis limited by the transducers speeds (Fig 1).

The greater part of research on optical computers has been done with photographic transparencies [1,2], used as data input, storage and output elements, which exclude real time processing . Since fast and sensitive photodetectors are now available, development of real time optical computing techniques depends upon progress in performances of input transducers (page composers) and of reusable photosensitive storage elements.

In this paper, we give a review of properties of solid state photosensitive devices and applications to optical processing.

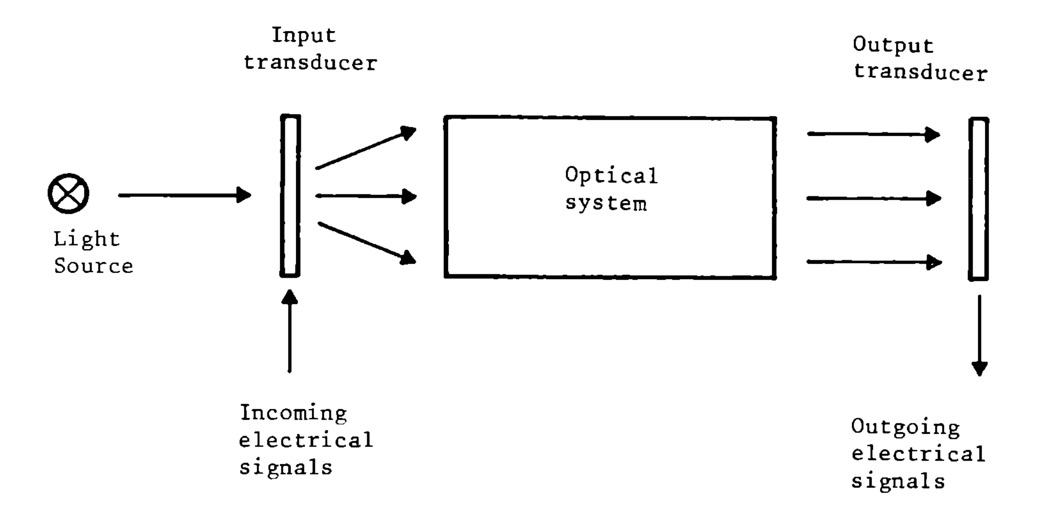


Figure 1 :
General block diagram of a typical coherent computer system.

Such devices are defined by the following properties :
a) Real time processing , i.e. reconstruction in situ of the recorded image without development, or with fast physical development by electrical, optical or thermal methods.
b) Contrast and photosensitivity controlled by combinations of optical and electrical parameters.
c) Storage time of the recorded image at least equal to the processing time of the computer.
d) Complete or selective erasure of the recorded data and rewriting capabilities

Such devices can be divided in two groups ; hybrid devices using combinations of photosensitive layers and electrooptic materials, and devices using materials both photosensitive and electrooptic.

II. PHOTOCONDUCTOR - ELECTROOPTIC DEVICES (PCEL)

The general principle of operation is shown on Fig 2. The transparency to be recorded is projected on the device, using the light source S1 and the lens L1, and produces a spatial change of photoconductor resistance r(x,y). When a voltage source V is connected to the transparent electrodes, the voltage applied to the electrooptic layer in the steady state is :

$$V(x,y) = V.\rho\left[\rho + r(x,y)\right]^{-1}$$

where ρ is the resistance of the electrooptic material.

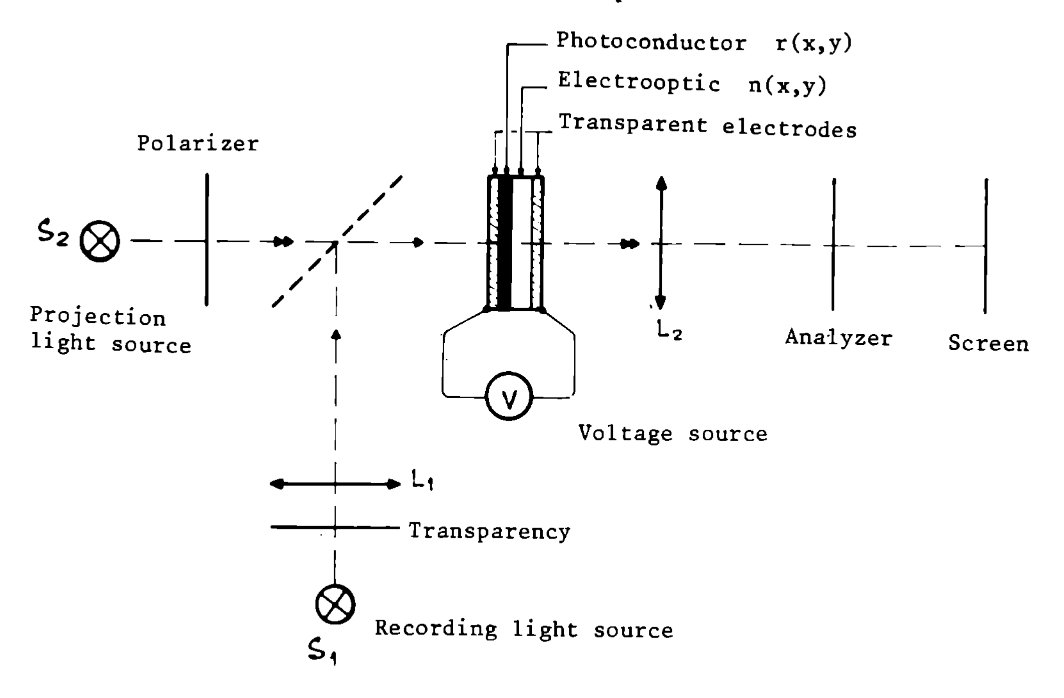


Figure 2 :
Photoconductor - Electrooptic device used in the transmission mode.

The spatial voltage distribution V(x,y) causes a spatial change in the refractive index of the electrooptic material ; this new phase image can be projected on a screen by the light source S_2 and L_2 as an amplitude image using crossed polarizers and analysers. This later operation is possible only when the photoconductor is insensitive to the wavelength of S_2.

As described, such device shows no memory capability and would have limited applications. Storage of the projected transparency can be obtained following two methods.
1) The spatial distribution V(x,y) can be stored electrostatically after the projection of the transparency during a time which depends on resistivities and dielectric constants of the photoconductor and the electrooptic material.
2) The spatial distribution δn(x,y) can be stored when the applied voltage causes in the electrooptic material a switching from a remanent state of refractive index to an other. This technique has been applied to transparent ferroelectric materials which shows two on more remanent values of the refractive index.

A) PCEL Using electrostatic storage

An example of such PCEL is the Phototitus [3]. The electrooptic material is a high resistivity KH_2PO_4 single crystal ; the

photoconductor is an amorphous Se layer. Operating temperature is - 50°C, which raises both the crystal resistivity and electro-optic coefficient. A dielectric layer, deposited between the Se layer and the KH_2PO_4 crystal allows reflection mode operation, which doubles the optical path in the crystal. Photosensitivity is 10 μJ/cm² at the recording wavelength λ = 420 nm ; image recording time is 10 μs. The spatial resolution is 70 lines per millimeter and remanence time 5 minutes.

The following analogical and digital operations can be performed :

- Sum of two images A and B is performed by recording B after A, with voltage V applied.
- Product of A and B is obtained by recording in one time the superimposed A and B transparencies. The corresponding Boolean operations are respectively O R and A N D.
- Recording image A with the applied voltage V and reading with applied voltage - V causes contrast reversal :
- The optical transfer function of the device is proportional to $\sin^2 kV$; therefore the voltage shift during reading causes the inversion in the slope of the transfer function, which in turn reverses the contrast of the stored image (Fig 3).
- Recording a first image A with applied voltage V and recording a second image B with applied voltage -V causes erasure of parts common to images A and B (addition of opposite voltages). This effect is used to perform image differences, and spatial differentiation. The corresponding Boolean operation on digital images is EXCLUSIVE OR.

Erasure is achieved by highly illuminating the crystal, with short circuited electrodes : this causes the complete flow of the stored charges.

B) PCEL Devices using ferroelectric storage (FEPC)

Three kinds of effects associated with electrical polarization switching in ferroelectric materials has been used in PCEL devices : remanent changes in birefringence, in light scattering angle and in surface deformation.

I) FEPC devices using changes in birefringence

In single crystals such as $Gd_2(M_0O_4)_3$ [4] Boracites [5] and $Bi_4Ti_3O_{12}$ [6], polarization reversal is accompanied by a rotation of the optical indicatrix from a stable position to an other (Fig 4a) . The difference δn in birefringences between these two states can be used in high contrast FEPC devices when the crystal thickness t is such as t δn = λ/2 (transmission mode) or t δn = λ/4 (transmission - reflection - transmission mode, similar to Phototitus operating mode).These two birefringence states allow digital recording only. Nevertheless, holograms have been recorded by Keneman et al in $Bi_4Ti_3O_{12}$, using a Zn Se photoconductive layer [7] ; sensitivity at λ = 488 nm is 1 mJ/cm² and resolution

800 lines/mm in a 25 μm thick crystal.

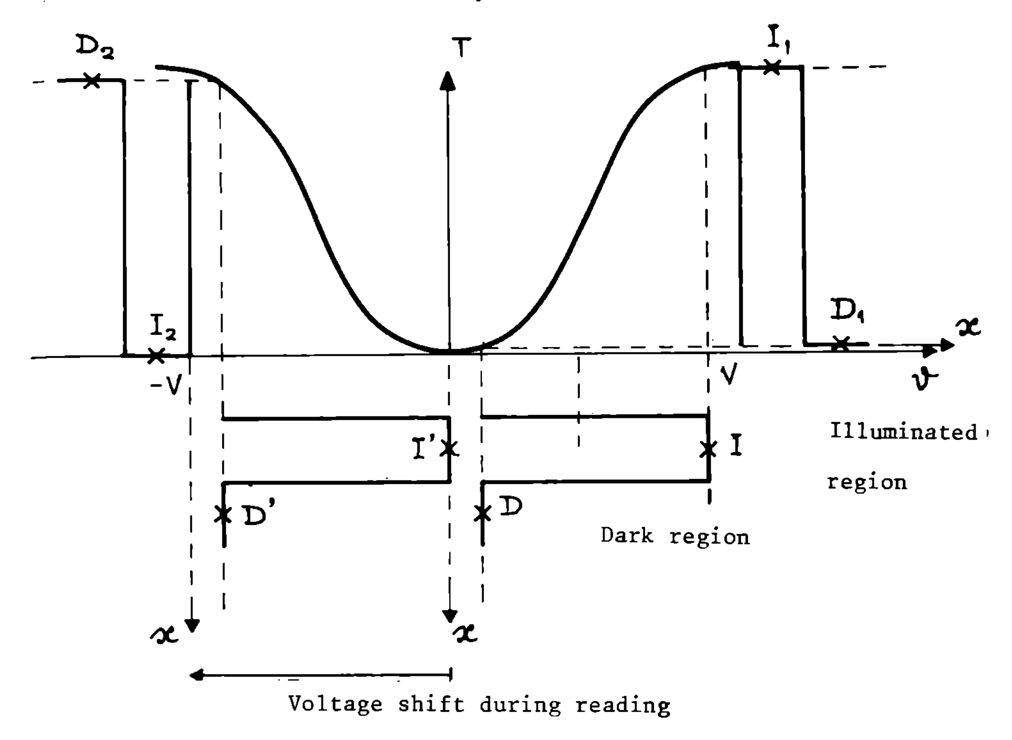


Figure 3 :
Light transmittance versus applied voltage and contrast reversal by voltage shift during reading.

In transparent ferroelectric ceramics such as $(Pb_{1-x}La_x)(Zr_{0,65}Ti_{0,35})O_3$ with $5.10^{-2} \leqslant x \leqslant 7.10^{-2}$, (x/65/35 PLZT), alignment of the individual grain polarization directions can be performed by applying either an electrical field either a mechanical stress. The relative number of aligned grains depends upon the amplitude of the external stresses; therefore the macroscopic birefringence characteristic of the sample depends both on amplitude and direction of the external stresses. The feasibility of a device using changes in birefringence of PLZT samples associates with PVK or CdS photoconductive layers has been first demontrated by Maldonado and Maitzler, in transmission and in reflection modes. [8.] . The PEPC device (FERPIC) is bonded on a plexiglass substrate, which is flexed and induces a permanent mechanical stress in the plane of the PLZT sample. The ferroelectric domains are aligned along the tensile axis in the plane of the sample as shown in Figure 4b, where is represented the optical indicatrix. The electrical polarization has no component in the perpendicular direction. Applying an electrical field through

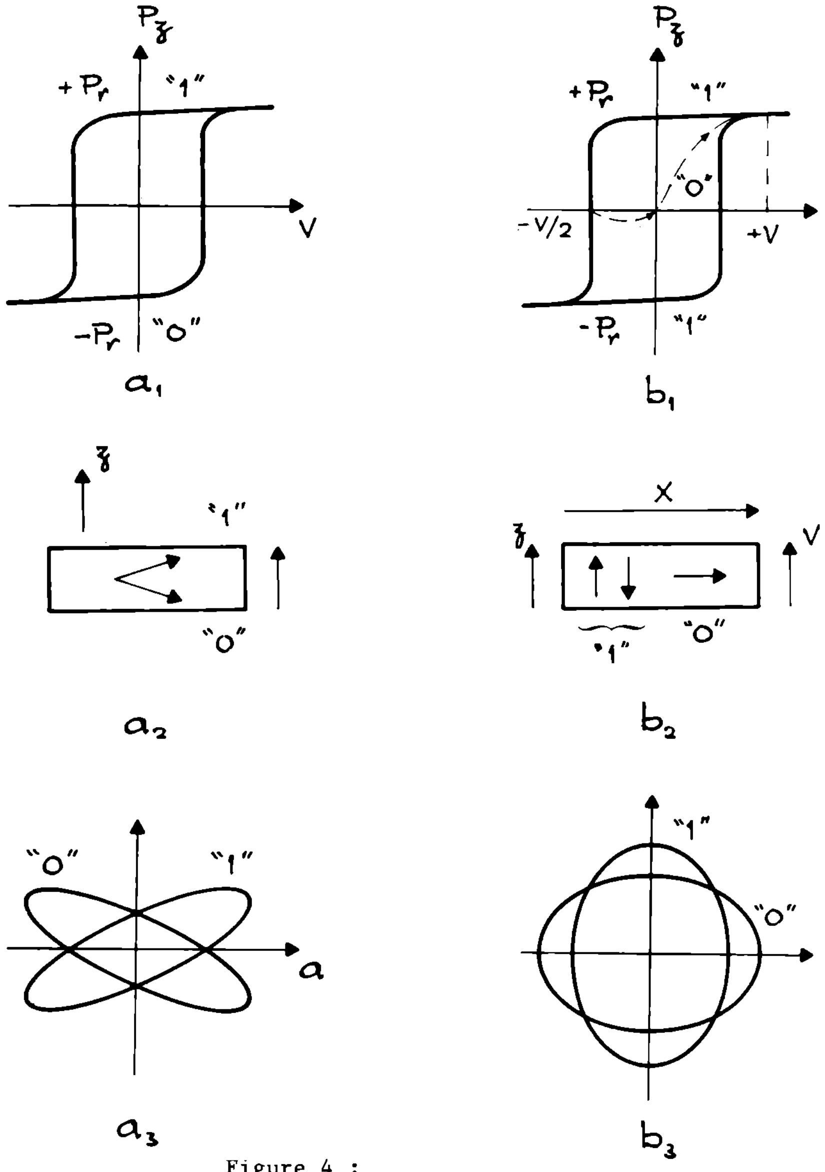


Figure 4 :
a_1, a_2, a_3 Hysteresis loop, polarization switching and optical indicatrix for $Bi_4Ti_3O_{12}$.
b_1, b_2, b_3 The same for strained PLZT.

the illuminated photoconductor causes a remanent alignment of the ferroelectric domains, in the direction of the field, and a 90° rotation of the optical indicatrix. The polarization component gets a non zero value P_z; the differences in birefringence between the two states $P_z = 0$ and $P_z \neq 0$ are used as previously described

for the spatial modulation of the reading light beam.

In a reflection mode device, using a 75 μm thick PLZT sample and PVK (Polyvinyl Carbazole), recording is achieved in 2s with a 200 mW/cm^2 light beam at λ = 576 nm. The sample is first uniformely illuminated with -V/2 applied (V = 70 volts), which ensures $P_z = 0$ ("0" state) ; the transparency to be recorded is projected with V applied : in the illuminated regions, polarization switches to the P_z value (" 1 " state).

II) FEPC devices using changes in scattering angle

In transparent polycristalline ferroelectrics such as PLZT samples, the macroscopic depolarized state corresponds to a random distribution of domains orientations. Each domain being birefringent by itself, their random orientations create refractive index inhomogeneities which cause scattering of an incident light beam. When the domains are oriented by an electric field for instance, a macroscopic birefringence is induced, and the scattering angle is decreased to a low value. Different remanent scattering angles can be induced in coarse grained PLZT ceramics to which correspond different optical transmittances, as shown in Fig 5 by Smith and Lamb [9].

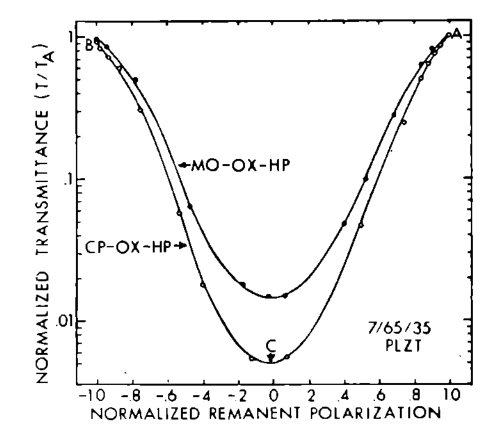


Figure 5 :
Normalized transmittance T/T_A vs normalized remanent polarization for two 7/65/35 PLZT plates 0.375 mm thick. Plates were initially poled to saturation remanence by applying 450 V dc (E=12 KV/cm). After Ref 9 .

The corresponding device, called CERAMPIC , does not require strain bias nor polarized light. Typical spatial resolution is 80 lines/mm.

An other scattering FEPC device was proposed by Kumada et al [10] in which the scattering state is a ferroelectric one (with

randomly oriented domains) and the non scattering one is an anti-ferroelectric one, the transition between the two states being electrically induced. The particular 7.6/70/30 PLZT composition was found which shows the typical double hysteresis loop (Fig 6 a) and a transmittance ratio between the two states of the order of 10^3 (Fig 6b).

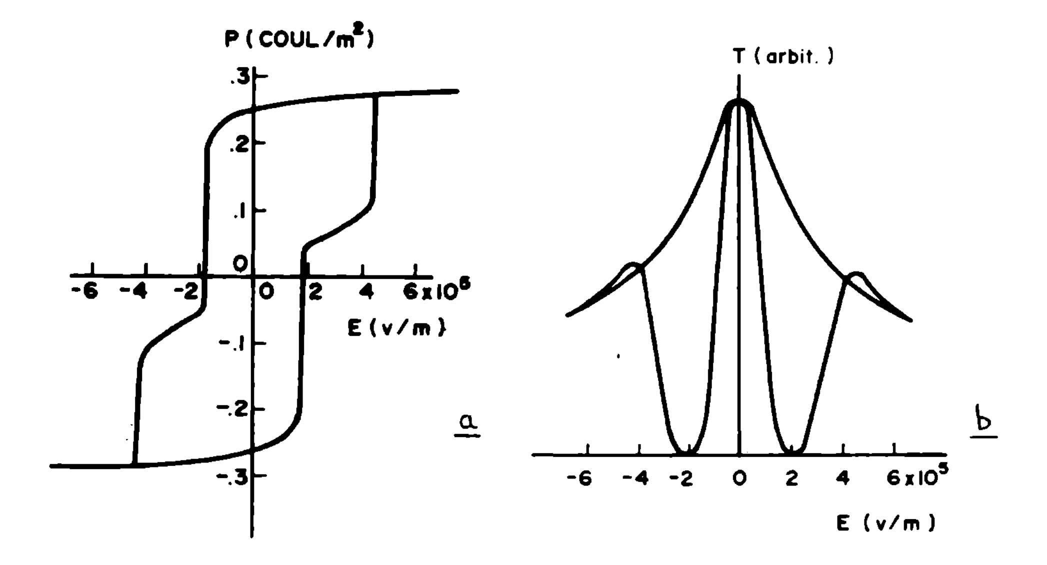


Figure 6 :
a) Double hysterisis loop of the Antiferroelectric-Ferroelectric 7.6/70/30 PLZT observed by a 5 mHz electric field at room temperature.
b) Field induced transmittance T-E curve of 7.6/70/30 PLZT . After Ref 10.

III) FEPC devices using changes in surface deformations

It has been noticed by Land and Smith [11] that localized domain switching in PLZT ceramics causes surface deformations which correspond to locally switched areas. Based on this observation, a FEPC device called FERICON was constructed, using reflection on the metallized surface deformed by local polarization switchings (Fig 7). Fine grained 7/65/35 PLZT ceramics were used, which showed 0.5 µm amplitude deformation in 250 µm thick samples, and produced 7 db contrast ratio for a 40 line pairs/mm resolution.

IV) Optical processing with FEPC Devices

The three kinds of FEPC devices allow to perform product of images (A N D) and addition (O R) since, in the later case, they are able to store the first image before recording the second. They have common capabilities such as performing inversion, selective erasure and difference (EXCLUSIVE OR). For instance , in 7/65/35 PLZT

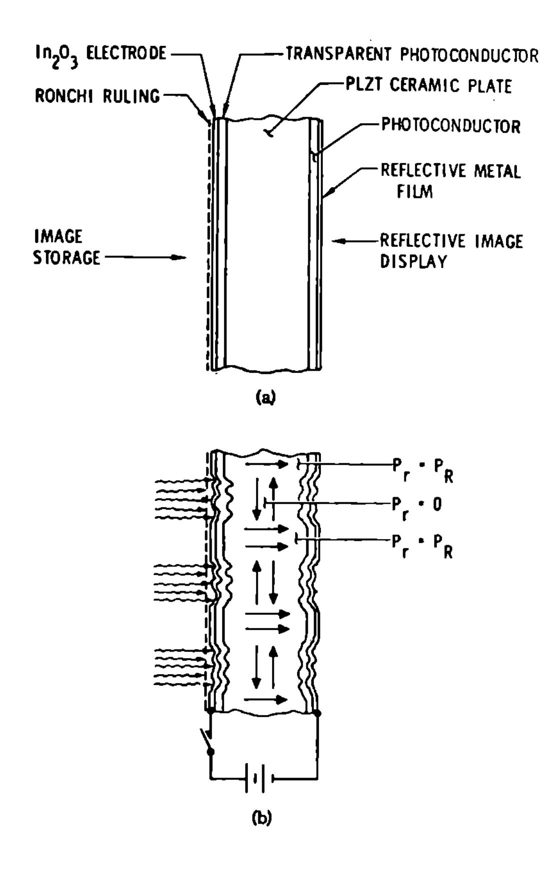


Figure 7 :
a) Fericon structure ; b) illustration of the surface deformations resulting from a bar pattern stored through the Ronchi ruling.After Ref II.

ceramics, assuming that the "1" state is associated with the non zero values of polarization $P = \pm Pr$ and the "0" state to $P = 0$, operations on images are obtained as follows : (Fig 4b)
-starting from the homogeneous state $P = 0$, the FEPC device is illuminated by the image A to be recorded, with a voltage V applied which switches the polarization to $P = Pr$ in the illuminated regions. The image is then recorded. Illuminating the FEPC device with the second image B; the voltage V being applied, causes the points "1" of image A illuminated by points "1" of image B to return in the "0" state. The truth table of such an operation corresponds to the EXCLUSIVE OR between digital image and difference between analogical images. When the image B is uniformly clear (B="1"), the former operation reverses the contrast of the recorded image A. Erasure is achieved by uniform illumination with V applied, which switches the FEPC in the "1" state, and a second uniform illumination with -V/2 applied, which turns off the device to the

uniform "0" state.

Although FEPC devices show capabilities of operations on images, the main applications would be rather in the field of displays, since they have the unique property of storage without energy consumption.

III. PHOTOSENSITIVE ELECTROOPTIC DEVICES

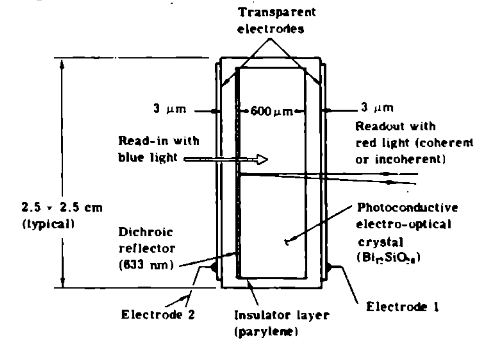
In these devices, one active material only ensures both the conversion of light spatial distribution into electric field distribution and the reading light modulation.

Photoelectrons excited during recording leave their origin centers by diffusion or drift under the influence of an applied field ; they can be trapped either by blocking electrodes by high capture cross section centers inside the material. In both cases, inhomogeneous illumination photoinduces a space charge field distribution which modulates the refractive index via the electrooptic effect.

Let us consider the two cases of trapping processes which are at the origin of two kinds of devices.

A) Photosensitive electrooptic materials without trapping centers

Materials such as ZnS, ZnSe, $Bi_{12}\ Ge\ O_{20}$ and $Bi_{12}\ Si\ O_{20}$ are both photoconductive and electrooptic. It was demonstrated by Feinlib et al [12,13] that a real time optical processor could be fabricated using the $Bi_{12}\ Si\ O_{20}$ crystal. Composition of the Pockels Readout Optical Modulator (PROM) is shown on Fig. 8.



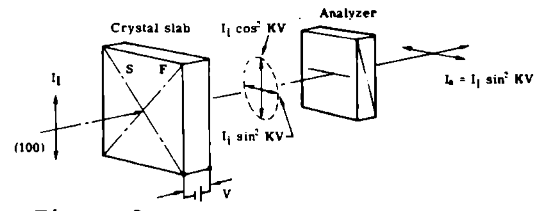


Figure 8 :
Composition and typical dimensions of PROM (After Ref. 12).

One face of the crystal is coated with a 633 nm dichroic filter (operation in the reflection mode) ; the crystal is encapsulated in a parylene insulator and transparent electrodes are deposited on the main surfaces of the device.

The sequence of image recording and reading is shown in Fig.9.

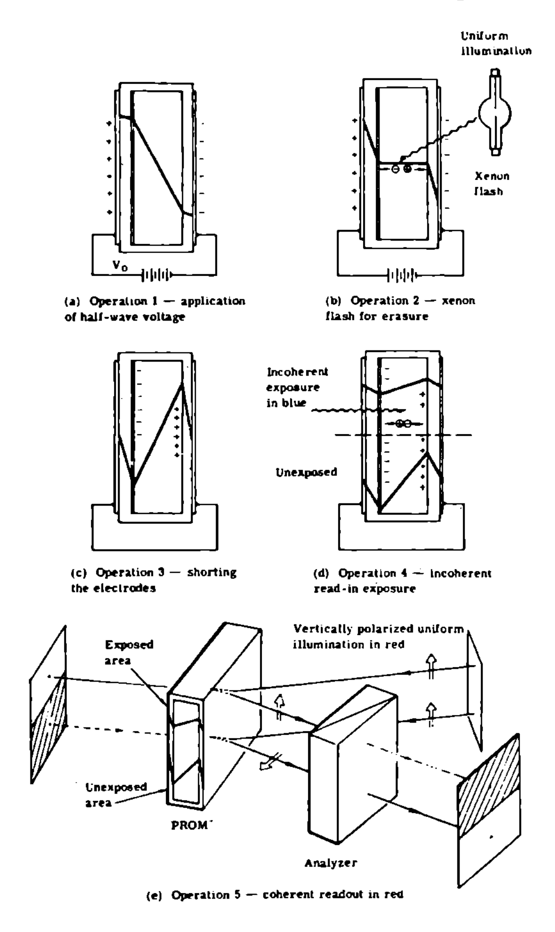


Figure 9 :
Operation of PROM (After Ref.I2).

After applying the half wave voltage (V=2.000 v), a uniform illumination is achieved using a xenon flash : excited photoelectrons are displaced by the applied field and stored on the insulator film. They produce a space charge field which cancels the applied field ; when short circuiting the electrodes the field is reversed in the crystal. The image to be recorded is projected on the PROM using blue light beam. In the brigth parts of the image, bits "1" of a binary transparency the field is cancelled again, and remains unchanged in the

dark regions. (bits "0") ; reading is achieved using polarized red light in reflection.

Contrast inversion is achieved by reapplying to the PROM the initial voltage V after the image is stored, since there is a quadratic relationship between output light intensity and voltage remaining in the crystal.

When recording a second image B after a first image A is stored, the spatial voltage distribution corresponding to A is selectively erased by the bright parts of image B. The Boolean operations performed are :

- NOR when recording B with no voltage applied.
- INHIBITION when recording B with V applied. Operating characteristics of PROM devices are typically a 300 line pairs/mm resolution, sensitivity of 10 $\mu J/cm^2$ using a He-Cd laser and a storage time of few minutes in the dark

B) Photosensitive electrooptic materials with trapping centers

Since Ashkin et al have observed [14] the laser induced change of refractive index in Li Nb O_3 (photorefractive effect), this effect was extensively studied for application to optical storage. As explained by Glass [15], the photorefractive effect is due to photocarriers displacements and trapping which induce local space charge fields and refractive index changes via the electrooptic effect.

Photorefractive effect is achieved in polarized materials only. Electrical polarization can be induced by an external field in materials with non polar crystalline lattices (ferroelectrics at temperature higher than their Curie temperatures) ; Single domain pyroelectric materials (for instance, ferroelectrics at temperature lower than their Curie temperatures) have a permanent polarization, so that changes in refractive index can be photoinduced without applying an external field.

Although it was fundamentally considered for applications to optical storage, the photorefractive effect can be used for optical processing, as shown in the two following examples.

I) Non polar photosensitive ferroelectrics

Example of non polar photosensitive ferroelectric is 9/65/35 PLTZ Ceramic [16]. In this material, the field dependence of birefringence is nearly quadratic, and follows with little hysteresis the relation

$$\Delta n\ (E) = A\ E^2$$

When illuminating a ceramic sample with a focused Argon light beam and simultaneously applying an electric field E_a, the field induced birefringence $\Delta n_a = \Delta n\ (E_a)$ is cancelled. This effect is interpreted by assuming that a screening field E_s has been photoinduced, with amplitude equal to E_a and opposite polarity ; therefore the local field in the illuminated region is $E\ell = E_a + E_s = 0$ and

$\Delta n = 0$. When the field E_a is removed, the birefringence in the illuminated regions becomes $\Delta n(E\ ) = \Delta n_a$. Using a scanned and modulated argon laser beam, this process allows optical recording ; the recorded data appear as bright dots on a black background between crossed polarizers.

Selective erasure of a recorded dot is obtained by illuminating this dot with the argon laser beam, without applying an external field : the local field $E\ell = E_a$ displaces the photoelectrons in a direction opposite to the recording direction so that the field in the erased dot is cancelled again (Fig 10).

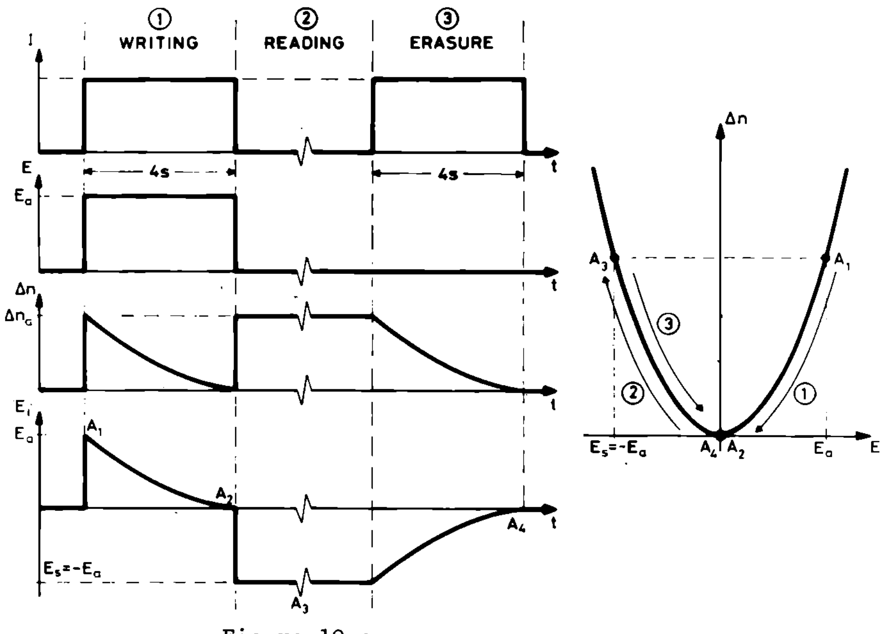


Figure 10 :
Recording, reading and erasure processes in 9/65/35 PLZT. I=2,20w cm^{-2} at λ = 488nm E_a =3,3 KV cm^{-1}, Δn=5.5 IO^{-4}.

When recording a set of points corresponding to an image A with the field E_a applied, and recording then a second image B with short circuited electrodes, the resulting stored image is given by the Boolean INHIBITION operation between data of images A and B. When two images C and D are recorded, the first with field E_a applied the second with the field $-E_a$ applied, the contrast of the two images can be reversed by changing the field polarity from E_a to $-E_a$; when reading with the fields $-E_a/2$ or $+E_a/2$ image C or image D can be selected (Fig 11).

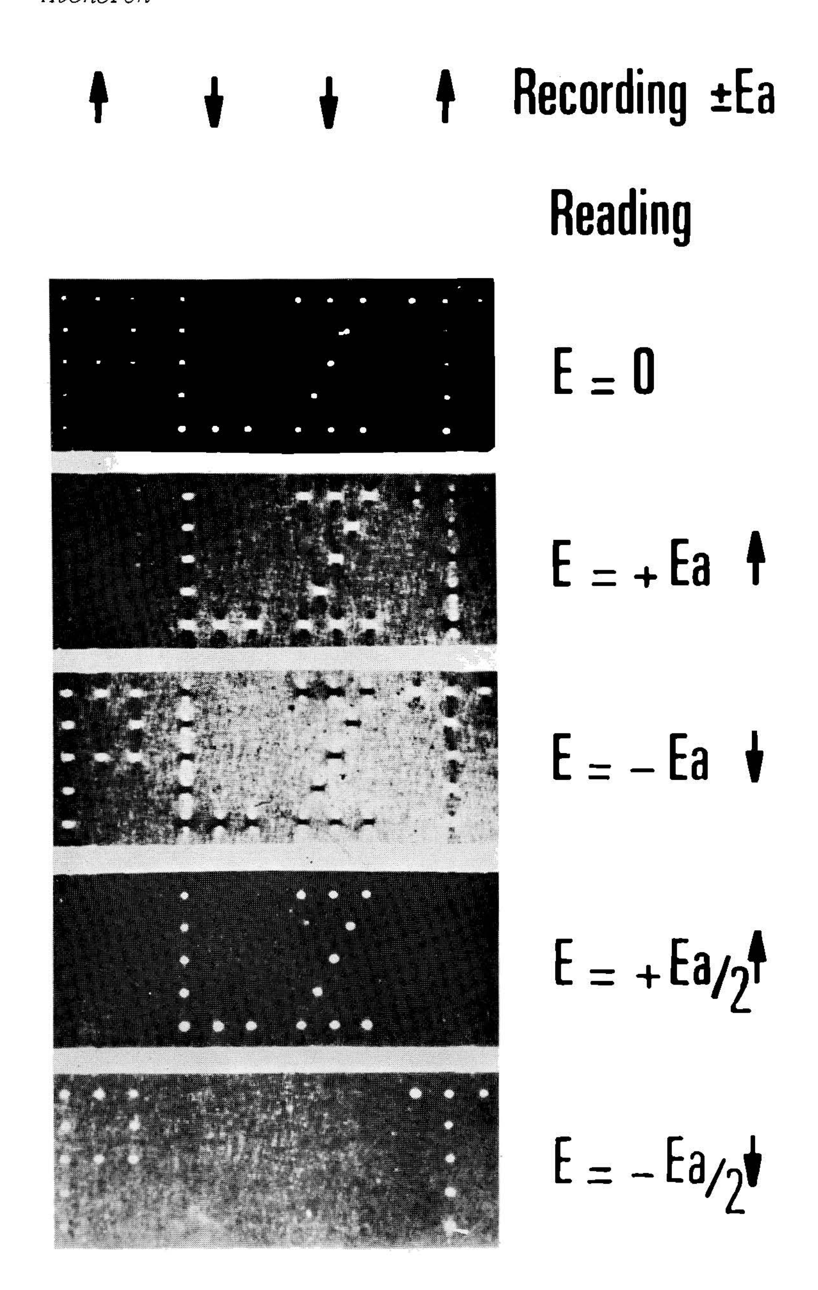


Figure 11 :
Images stored in 9/65/35 PLZT (100 μm thick) observed between crossed polarizers for different applied fields during reading. Arrows represent the field directions during recording and reading.

PLZT ceramics show low photosensitivity since light energy density for cancelling the field induced birefringence $\Delta n = 5.5\ 10^{-4}$ is typically I= 880 w/cm^2 . This photosensitivity would be raised by suitable doping.

Spatial resolution is high enough, since holograms with 10 µm fringe spacing can be recorded, with diffraction efficiencies larger than 30% in 100 µm thick samples [17] .

II) Polar photosensitive ferroelectrics

Polar photosensitive ferroelectrics such as iron doped Li Nb O_3 crystals are suitable for optical recording of stacked Fourier holograms, since light energy density required for recording is lower than for optical erasure (asymetrical read-write cycle) as shown in Fig 13 a. Symetrical cycles can be achieved using a coherent erasure technique : the hologram to be erased is recorded again with a π phase shift introduced in the reference beam (Fig 12, 13 b) ; this second recording produces the complementary refractive index spatial modulation, and therefore cancels the first recorded modulation.

Experimental setup

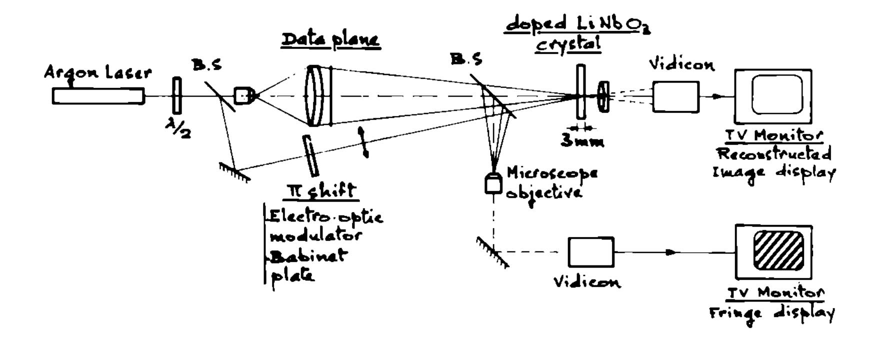


Figure 12 :
Experimental holographic Set up, with adjustable phase shift on the reference beam, measured by fringes display.

This is the first known technique which provides selective erasure of holograms superimposed in the same volume (Fig 14 a).

Owing to the linearity of Fourier transform we have shown that recording the holograms of a first image A then the hologram of a second image B with a π shift in the reference beam causes erasure in the reconstructed image of the common parts of A and B (Fig 14b).

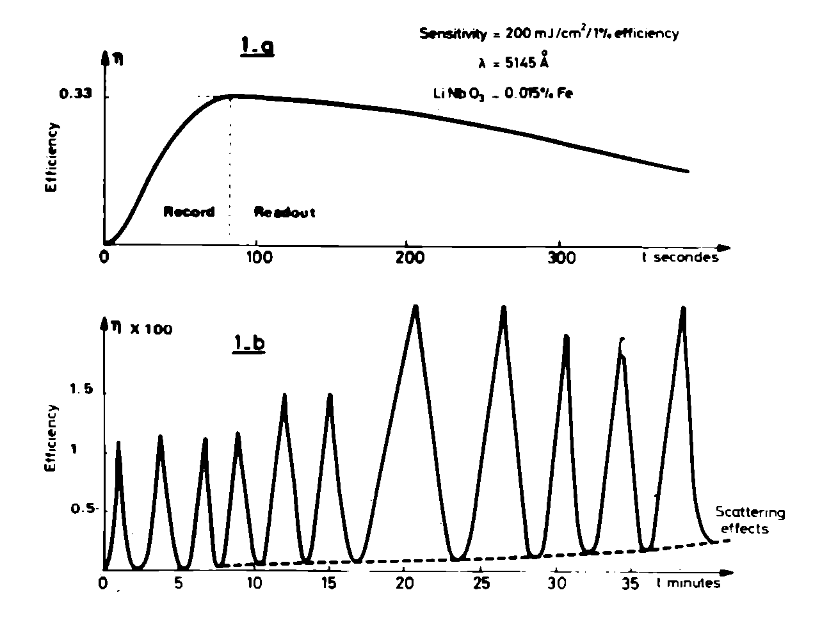
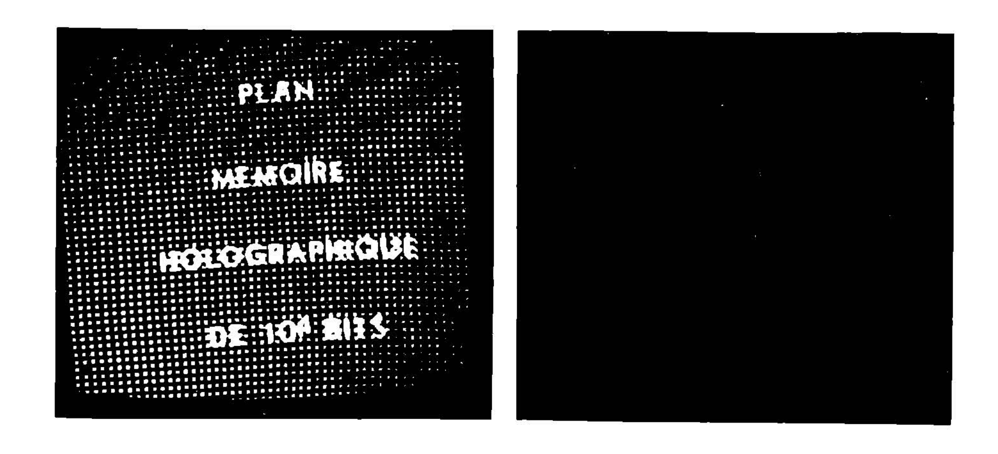


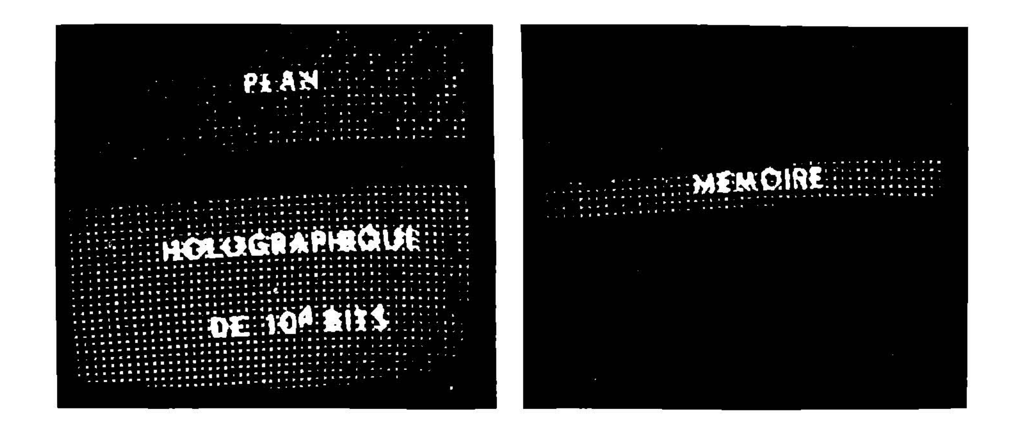
Figure 13 :
a) *Holographic diffraction efficiency versus time in reduced Li Nb O_3 0.015% Fe (3mm thick). Asymetrical Recording-erasure cycle.*
b) *Symetrical Recording - Coherent erasure cycles in the same crystal.*

This operation on digital images is the Boolean EXCLUSIVE OR. In the particular case where B=1, i.e. when B is a uniform transparency, this operation reverses the contrast of image A (Fig 15).

These coherent erasure techniques look promising for optical processing on three dimensional holograms recorded in photosensitive ferroelectrics ; they are mainly limited by high light energy required for recording (0,1 to 1 J/cm^2 at λ = 488 nm) and by light scattering induced during permanent reading or by sequences of write-read-erase cycles.

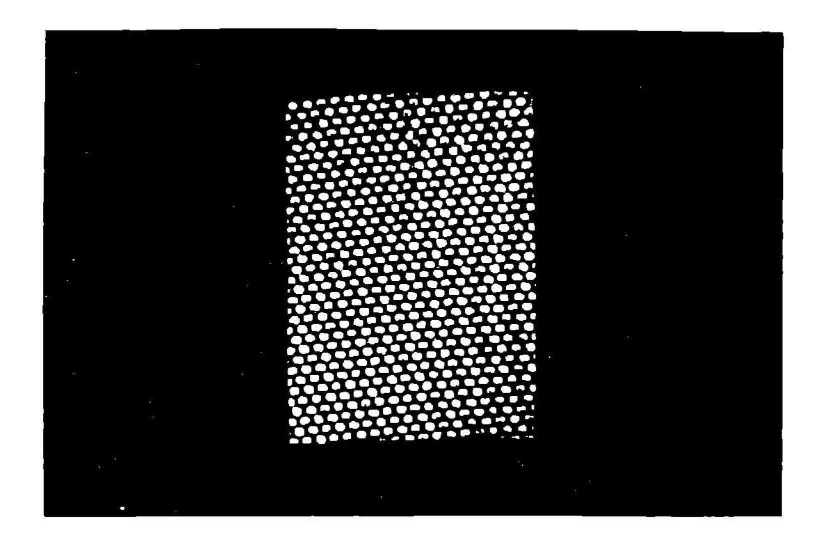


.a

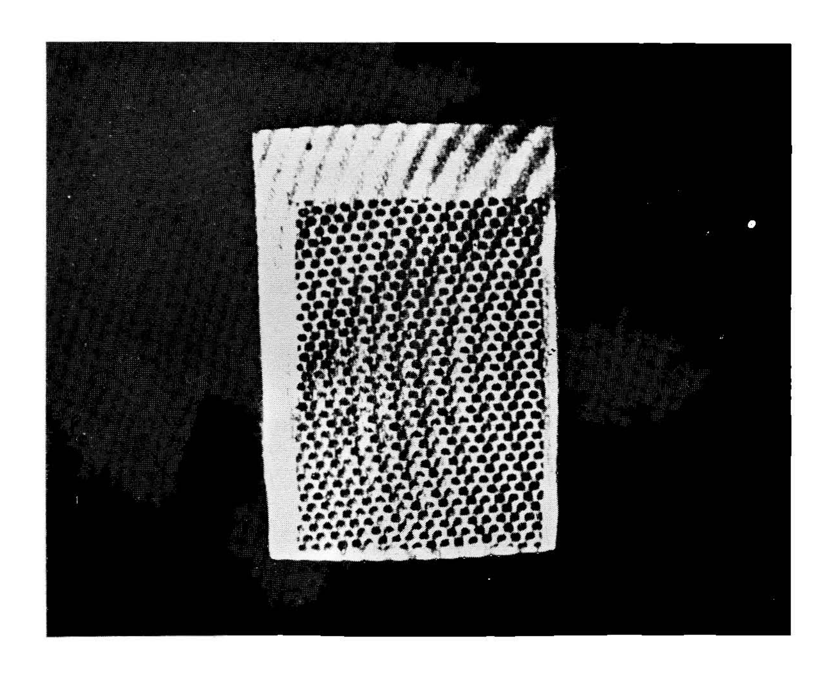


.b

Figure 14 :
a) Reconstructed image and erasure
b) Selective coherent erasure in a page.



-a



-b

Figure 15 :
a) Original image reconstruction
b) Image reconstruction with reversed contrast.

CONCLUSION

In the hybrid photoconductor-electrooptic devices conversion of recording light distribution into electric field distribution and modulation of the reading light beam are performed by separate materials, which, in some degree, can be independentely optimized for a given application ; spatial resolution is limited by the thickness of the electrooptic material. In photo-sensitive electrooptic devices the two functions are strongly coupled in $Bi_{12} Si O_{20}$ for instance, but weakly coupled in photosensitive ferroelectrics such as $Li Nb O_3$ crystals : photosensitization by doping has little effect on their electrooptic coefficients.

Although the photorefractive process needs recording light energy densities several orders of magnitude larger than the photoconductor-electrooptic combination, it shows the unique property of optical processing whith holograms stacked in the volume of the crystals, and selective erasure of three dimensional holograms.

In anay case, processes and devices we have reviewed are still undeveloped and require for practical applications extensive research work in the fields of optical materials synthesis and knowledge.

REFERENCES

[1] A.R. SHULMAN
Optical data processing . John Wiley and sons 1970.

[2] K. PRESTON Jr
Coherent Optical Computers . Mc Graw Hill 1972.

[3] M. GRENOT , J. PERGALE, J. DONJON and G. MARIE
Appl. Phys. Lett. 21 , 83 (1972)

[4] A. KUMADA
Ferroelectrics 3 , 115 1972.

[5] J. KOBAYASHI , N. YAMADA, H. SCHMID and H. SCHACHNER
Phys. Rev. B, 1 , 3801 1970.

[6] S.E. CUMMINS and L.E. CROSS
J. Appl. Phys. 39 , 2268 1968.

[7] S.A. KENNEMAN, A. MILLER and G.W. TAYLOR
Appl. Opt. 9 , 2279 1970.

[8] J.R. MALDONADO and A.H. MEITZLER
Proc. I.E.E.E. 59 , 368 1971.

[9] W.D. SMITH and C.E. LAND
Appl. Phys. Lett. 20 , 169 1972.

[10] A. KUMADA , G. TODA and Y. OTOMO
Proc. 5th Conf. Sol. State Devices TOKYO 1973
Suppl. Jour. Jap. Soc. Appl. Phys. 43 150 1974.

[11] C.E. LAND and W.D. SMITH .
Appl. Phys. Lett. 23 , 57 1973.

[12] J. FEINLIB
Laser Focus 43 Sept. 1973

[13] J. FEINLIB and D.S. OLIVER
Appl. Optics 11 , 2752 1972.

[14] A. ASHKIN , G.D. BOYD, J.M. DIEDZIC, R.G. SMITH, A.A. BALMAN
J.J. LEVINSTEIN and K. NASSAU
Appl. Phys. Lett. 9 72, 1966.

[15] A.M. GLASS
In this issue, p.

[16] F. MICHERON, J.M. ROUCHON , M. VERGNOLLE
Appl. Phys. Lett. 24 , 605 1974.

[17] F. MICHERON , A. HERMOSIN , G. BISMUTH, J. NICOLAS
C.R. Acad. Sci. B 274 - 361 1972.

APPLICATION OF TITUS AND PHOTOTITUS TO IMAGE AND DATA PROCESSING

J.P. Hazan
Laboratoire d'Electronique et de Physique Appliquée
94450 Limeil-Brévannes (France)

▪ I. INTRODUCTION

Titus [1] and Phototitus [2] are electro optical light valves with grey scale storage capabilities allowing real-time high-frame rate data processing in non-coherent as well as coherent light.

These devices use the Pockels effect in a KD_2PO_4 (DKDP) crystal, operated in the paraelectric phase and cooled down just above the Curie temperature. The cooling is effected by Peltier cells incorporated in the device.

The two devices differ in particular by the way the electrical charges are deposited on the crystal and by their addressing mode : electron beam addressing in the case of Titus and optical addressing by means of a photoconductor for Phototitus.

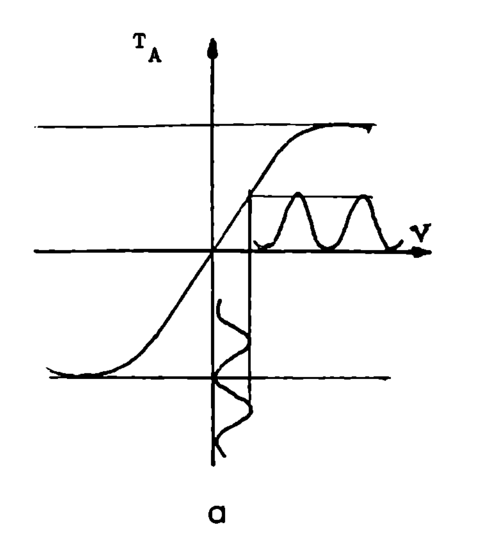
These devices have already been discussed in the literature [3,7], in this short paper we will first review relatively briefly their mode of operation, their characteristics and their performance and will stress more particularly a few of their applications to real-time data processing in coherent and non-coherent light.

▪ II. OPERATION OF TITUS AND PHOTOTITUS

A) General considerations

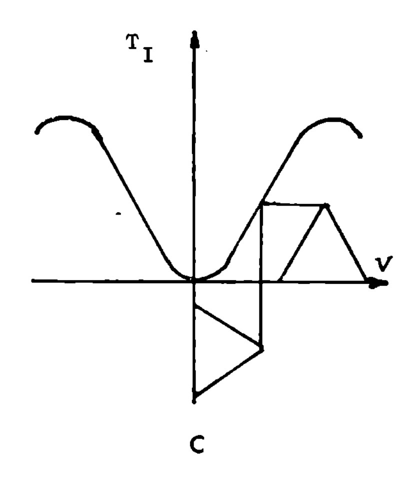
Before describing the mode of operation of these devices, we will discuss a few interesting features of the Pockels effect. The latter is a linear electro-optic effect occuring in certain crystals whereby a linearly polarized light is made elliptical by electrically induced birefringence. The modulated light can then be "analyzed" by means of a polarizer. Two characteristics are worth pointing out :

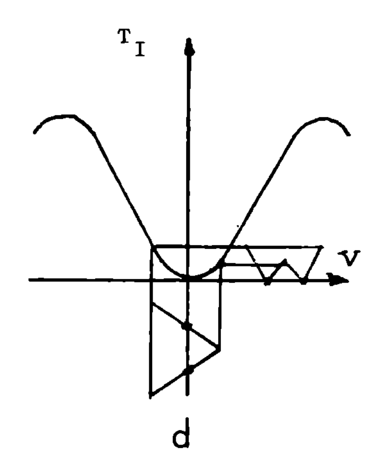
★ The modulation can be obtained with an electric field parallel to the direction of propagation of light (longitudinal effect). Thus, each point of the cross section of a polarized in-



b

$T_A = \sin kV$





$T_I = \sin^2 kV$

Figure 1 :
Modulation characteristic curves with respect to amplitude transmission (Fig.1a and 1b) and to intensity transmission (Fig 1c and 1d). The use of an electrical or optical bias allows :

- *the attenuation of the 0 order diffraction term by cancelling the average "DC component" (Fig 1a and 1b) ;*
- *the generation of equiluminance contours : a given grey level will be displayed in black (Fig 1c and 1d).*

cident light beam can be independently modulated by locally deposited electric charges on the surface of the crystal plate.

★ A second useful feature of the Pockels effect is its linearity ; indeed the phase shift ϕ between the light components parallel to the induced crystallographic axis will be proportional to the electric field E and the voltage V accross the crystal : $\phi \propto 2$ kV. This linearity will be of interest for example in the case of addition and subtraction of a uniform background or a grey scale image to another grey scale image.

If one "analyzes" the modulated light by means of a crossed polarizer, the amplitude and intensity transmittance T_A and T_I are given by :

$$T_A \propto \sin \frac{\phi}{2} = \sin kV$$

$$T_I = T_A^2 \propto \sin^2 \frac{\phi}{2} = \sin^2 kV$$

the corresponding modulation characteristic curves are shown on figure 1.

In the conventional mode of operation the modulating voltage is for example positive and varies from 0 to V_{max}. One can however easily shift the 0 level by subtracting a uniform phase shift for example by means of a DC electrical bias (Titus) or an optical bias placed in the read-out beam such as a Bravais compensator or an orientable quarter wave plate [4] or by subtracting a uniform background (Phototitus).

Thus, "visual" positive to negative conversion can easily be obtained by subtracting a uniform phase shift ϕ of the order of π. The possibility of continuous adjustment of the 0 level can lead to several interesting applications such as :

- contrast enhancement ;
- the generation, in black, of equiluminance contours while the picture is in memory (as explained in fig. 1c and 1d);
- the suppression or attenuation of the 0 order diffraction term (the "undiffracted" light) in coherent optical Fourier processing by cancelling the "DC" optical component of a picture [8] (fig. 1a and 1b).

As mentioned in the introduction the DKDP is operated just above its Curie temperature . Compared to ambient temperature operation, this allows a twelve-fold reduction of the voltage required for a given phase shift. The two devices are operated in the reflection read-out mode hence a further two-fold reduction ; thus 125 V accross the crystal are sufficient to yield a modulation efficiency of 75 % of the intensity of an incident polarized light-beam.

Another advantage of Curie point operation is the memory capabilities imparted to the devices. The intrinsic discharge time constant of the DKDP goes from 0.2 second at ambient temperature to about an hour at Curie temperature. In a practical device other factors will affect the memory, for Phototitus, for example, the photoconductor's role is important : in the dark the memory is of the order of half an hour whereas with an orange light of sufficient intensity for small screen (13 × 18 cm) projection, it is of several minutes ; for Titus, in the case of while light large screen projection, the memory is typically of the order of 10 to 15 minutes.

The cooling is effected by Peltier coolers incorporated in the devices.

B) Device operation and characteristics

I) Titus

Figure 2 illustrates the mode of operation of Titus. The video signal is applied between one side of the crystal and a grid placed close to the other side. The addressing is effected independently by a constant current electron-beam. The latter acts essentially as a flying short-circuit between the grid and the point of impact of the crystal. The various points of the target thus take the potential the video signal has at the moment they are hit by the beam, irrespective of their previous potential. One thus obtains an

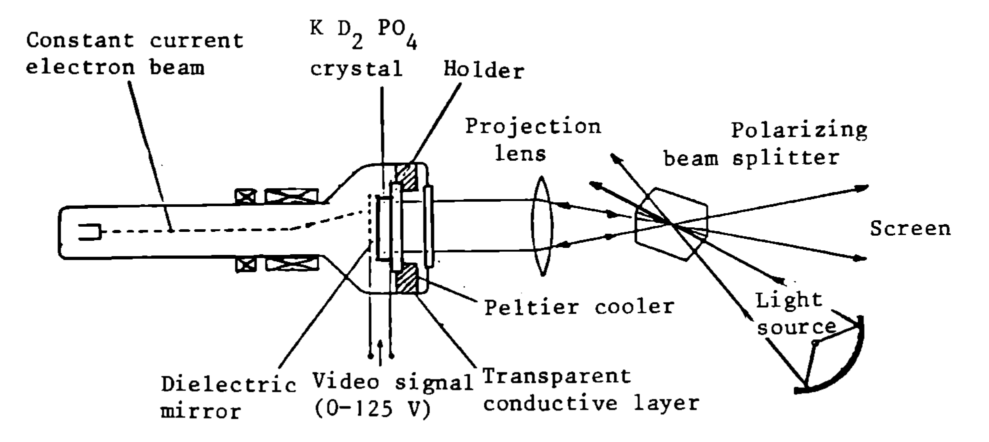


Figure 2 :
Principle of operation of Titus.
The video signal is applied between one side of the crystal and a grid placed close to the other side. The scanning is effected independently by a constant current electron beam. The input light beam is polarized and after reflection on the back side of the target, the modulated beam is analyzed by means of a crossed polarizer (a polarizing beam-splitter can be used as two crossed polarizers).

"electronic image" of the video signal. The input light beam is polarized for example by means of a calcite polarizing beam-splitter and after reflection on the back side of the target the modulated beam is analyzed by means of a crossed polarizer (the same polarizing beam splitter can also act as the analyzer).

It can be noted here that the light source being external it can be chosen in particular to be a powerful light source.

Indeed Titus has been primarily studied for large screen TV projection. We will not elaborate here on this broad field of application [1,3,4]. We will just mention some of its present characteristics :

- resolution : 750 elements/line (625 line standard, 25 frames/second);
- contrast : 30 to 50 ;
- output flux : 2 500 lm allowing $25m^2$ screen projection.

II) Phototitus

Phototitus (fig. 3) incorporates a DKDP target similar to that used in Titus but with an amorphous selenium photoconductor layer and a transparent conducting film deposited over the dielectric mirror.

In operation, an image is projected or scanned on the selenium layer while an electric field is applied accross the device ; this results in the transport of electric charges toward the DKDP crystal.

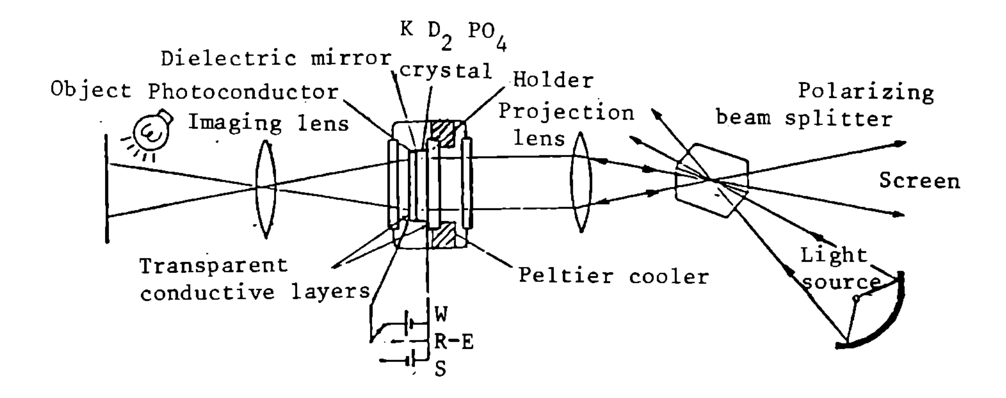


Figure 3 :
Principle of operation of phototitus
The input information is projected on the photoconductor while a voltage is applied across the device ; subtraction can be effected by inverting the polarity during the second write-in ; no external voltage is applied during reading or erasure. The latter is effected by flooding the photoconductor with light. As for Titus the read-out is effected in the reflection mode.

As selenium is an ambipolar photoconductor, positive as well as negative charges can be deposited on the addressing side of the DKDP crystal, depending on the sign of the applied voltage. The optical readout is effected in the reflection mode on the dielectric mirror which optically separates the writing and the reading functions. A light-blocking layer could also be used to screen further the photoconductor from the reading light. As with Titus, a polarizing beam splitter can be used to polarize the incident light and analyze the modulated light. During the reading and erasure cycles the electrodes are short-circuited; in the latter case the photoconductor is simultaneously flooded with light.

Phototitus is only suited for small to medium size screen projection. The device is still under development and its present characteristics can still be improved ; we can mention however :

- limit resolution : 5700 TV elements/line (75 l.p./mm) ;
- dimensions : 38×28 mm^2 target ;
- maximum contrast ratio*): >70/1 ;
- sensitivity : 10 $\mu J/cm^2$ (30/1 contrast) ;
- write-in time (whole image simultaneously) : $\sim$10 μs ;
- erasure time : $\sim$1000 μs (selective erasure possible).

III. APPLICATIONS TO IMAGE AND DATA PROCESSING

A) General considerations

The expression "data processing" immediately calls to the mind the words electronic, and more specifically digital, computer for which a tremendous amount of work has been expended.

Since the advent of lasers optical data processing has however stimulated a certain interest.

When one compares optical and digital processing one realizes that the two methods are in fact quite complementary ; the strong points of one method being usually found where the weak points of the other lie. Optical processing is less precise and somewhat less flexible than digital processing but it can perform multichannel, multiplex, complex processing much faster. The complementarity even extends to the type of mathematical operation which can be easily performed : digital computers will readily compute

---

*) The contrast is limited in particular by the natural birefingence of the DKDP. The contrast quoted is for full white/full black pictures and for incident angles of the order of $\sim 9°$ ($\sim 15°$ with a quartz compensating plate). In the case of coherent processing the incident beams are in general well collimated and much higher contrast can be obtained ; for full white/full black pictures it can exceed 1000/1.

additions, subtractions, multiplications and divisions ; integration is somewhat more lengthly ; however operations such as Fourier transform, correlation and convolution are appreciably more tedious to perform especially on two dimensional images. With optical processors it is on the contrary two dimensional Fourier transforms, correlation and convolution which are easiest to perform (in coherent light) ; to a certain extent integration and multiplication are relatively simple ; however division is very difficult and until the advent of devices such as Phototitus, real-time addition and subtraction of two images were quite difficult. Mr. Micheron has discussed on the other hand the possibility of performing various logic operations with certain type of devices.

Optical processing should therefore have its place in data handling, either alone or combined with electronic processing. Electro-optic light valves would then be a crucial component, especially if they can be used as real-time "incoherent-to-coherent convertors". In the sequel of the paper, we would like to illustrate, by a few examples, a few of the possible applications of Titus and Phototitus to non-coherent and coherent real-time optical processing.

In all these applications the initial, intermediate or resulting data or images can be observed at length, thanks to the grey scale memory capabilities of the devices.

B) Non-coherent light applications [4,6]

In non-coherent light applications, advantage is taken of the memory capabilities and of the possibilities of addition, subtraction and differentiation of grey scale images *) (Phototitus) as well as 0 level adjustment (Titus and Phototitus). On the other hand, Phototitus can also be used for ultra-violet, infra-red or X-ray to visible conversion.

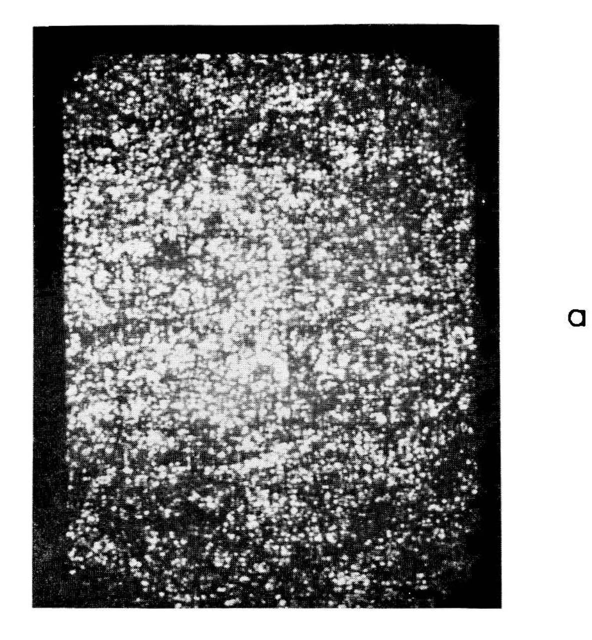
Let us now illustrate a few of these possibilities.

a) Addition :

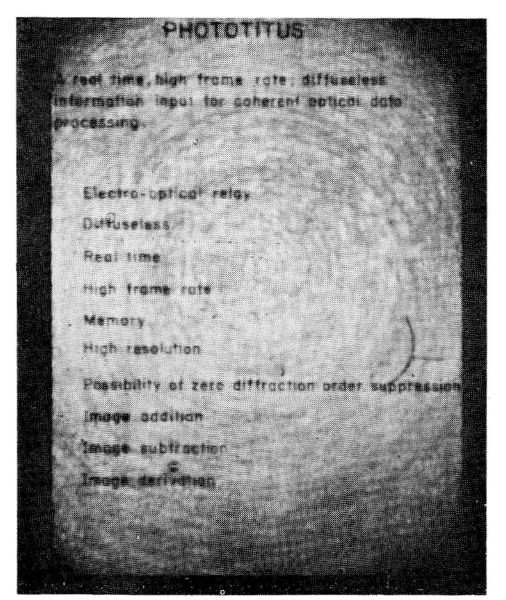
Addition is the basis of operations such as temporal and spatial integration.

Temporal integration can be used to increase the signal to noise ratio of pictures affected by uncorrelated noise. Figure 4 illus-

*) In its present version Titus does not allow addition and subtraction of images ; however, in a modified version, this would be quite feasible.



a



b

Figure 4 :
Signal to noise enhancement by temporal integration, on Phototitus, of pictures containing the same signal but affected by uncorrelated noise. Fig 4a is the image obtained by placing a fixed diffusor in front of a text. To obtain fig. 4b the diffusor was rotated during the exposure to simulate addition with uncorrelated noise (same total exposure as in fig. 4a).

trates this possibility. Figure 4a is the image obtained by placing a fixed diffusor in front of a text and projecting it on Phototitus. To obtain figure 4b, the diffusor was rotated during the exposure to simulate the addition of uncorrelated noise (the total exposure is the same as in figure 4a). The improvement can be observed in real-time ; the addition can be stopped and resumed at will ; the contrast of the final image can be improved by subtracting a uniform background by adjusting for example an optical bias.

Spatial integration can be used to enhance the fixed parts in a sum of several pictures taken e.g. from various points of view. This technique is useful in X-ray tomoscopy for which no focusing optics exist ; it allows to observe here in real-time, a given plane with good contrast while the other planes are faint and smeared out over the surface of the photoreceptor.

b) Subtraction :

Subtraction is the basis of a great number of image processing techniques ranging from contrast enhancement and equiluminance contours to change detection, image differentiation and small detail enhancement. One can distinguish background subtraction and image subtraction.

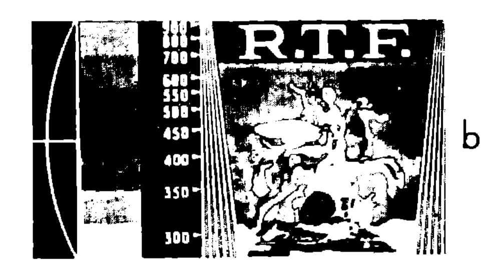
α) Background subtraction :

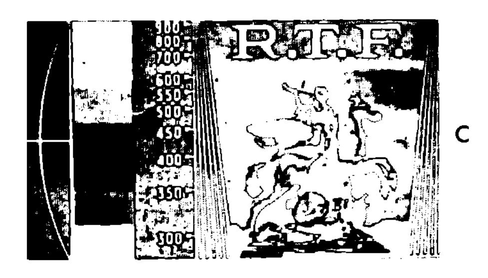
Background subtraction can be effected in a number of ways as discussed in section II-a by using for example an electrical bias (Titus), an optical bias (Titus and Phototitus) or by prorecording a negatively charged uniform image (Phototitus).

Background subtraction can be used for "virtual" positive to negative conversion★) and contrast enhancement. Another type of applications is the display of equiluminance contours. This stems from the modulation characteristics of the device and is explained on figure 1a. By subtracting a uniform background a particular grey level will be displayed in black, the neighbouring levels will of course be displayed in dark grey.

★) The term "visual" is used here to distinguish from electrical positive to negative conversion. Due to the quadratic nature of optical receivers, the intensity transmittance ($T_I = \sin^2 \frac{\phi}{2} = \sin^2 kV$) is an even function of V and is thus independant of the sign of V. A visual negative is observed by subtracting a phase shift of the order of π.







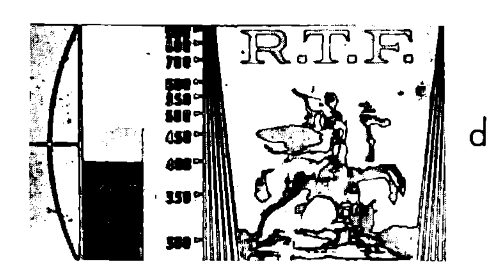


Figure 5 :
Generation of equiluminance contours by adjusting an optical bias with Phototitus. Fig. 5a : original picture ; Fig. 5b,c,d correspond to different bias settings while the image is in memory.

Figure 5 illustrates this possibility. Figure 5a is the original picture, figure 5b, c,d correspond to different optical biases. It is to be noted that since the image is in memory various grey levels can be analyzed successively in real-time ; one can thus make a coarse densitometric study of a picture ; another, related possibility, of background subtraction is the enhancement of grey features on a light background as shown in figure 6. One can also point out that in the case of a bright object on a dark background (or vice-versa), the proper choice of equiluminance grey will often yield a sort of "line drawing" contouring most of the object, in particular when the transition from dark to light is not very sharp ; one can even then analyze in some detail this transition.

β) Image subtraction

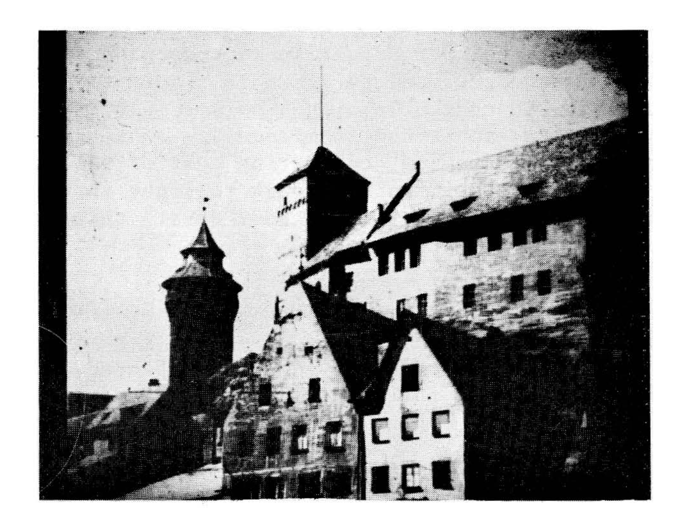
Due, in particular, to the ambipolar nature of the photoconductor used in Phototitus and to the memory effect of the device, grey scale image subtraction can readily be effected ★) as shown on figure 7. Figure 7a is the image displayed on Phototitus after recording a picture with a given electrical bias polarity. After this initial step, the same picture was reprojected again on the photoconductor but with the reverse polarity ; the resulting subtraction is shown on figure 7b. Real-time grey scale image subtraction can have a number of interesting applications such as change detection, image differentiation, small detail and contour enhancement, etc.

Figure 8 shows an example of change detection. A first scene was projected on the addressing side of Phototitus (Fig 8a), an object was displayed Fig 8b) and the resulting scene was subtracted from the first image recorded to yield figure 8c. As the photographic film is a quadratic detector, both removed and added objects appear in white on a dark background. If however , one adds an optical bias the removed object will appear in white and the added object in black on a grey background (Fig 8d) (the initial image was not totally subtracted to show the background setting). Of course once the ambiguity has been removed, the optical bias can be cancelled for higher contrast observation.

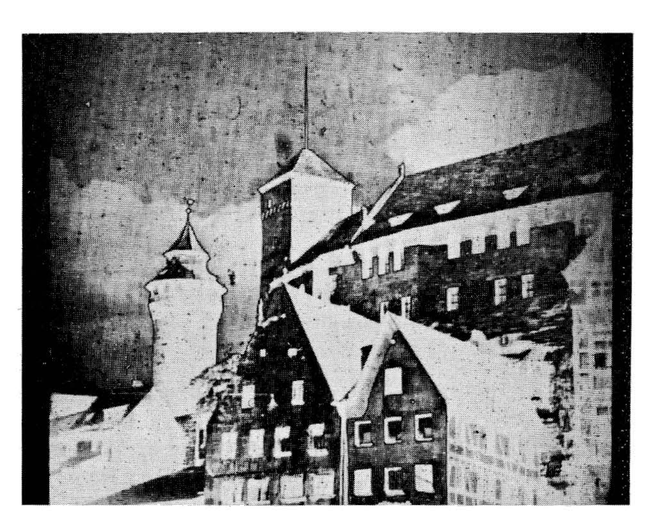
First and second order real-time image differentiation can be obtained with Phototitus. First order differentiation can be performed by recording a scene and subtracting the same scene slightly shifted in a given direction. In effect one thus performs

---

★) Titus, in a modified version, could also perform subtraction operations.







b

Figure 6 :
Enhancement of grey features on a light background (and of displayed dark features on dark background). Note also the line drawing-type contouring (see text).



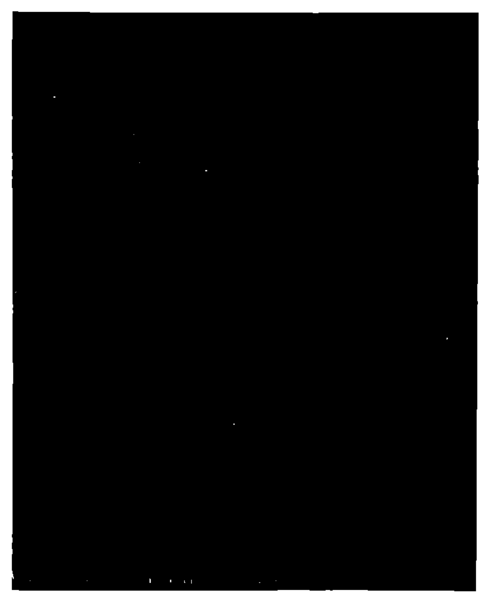
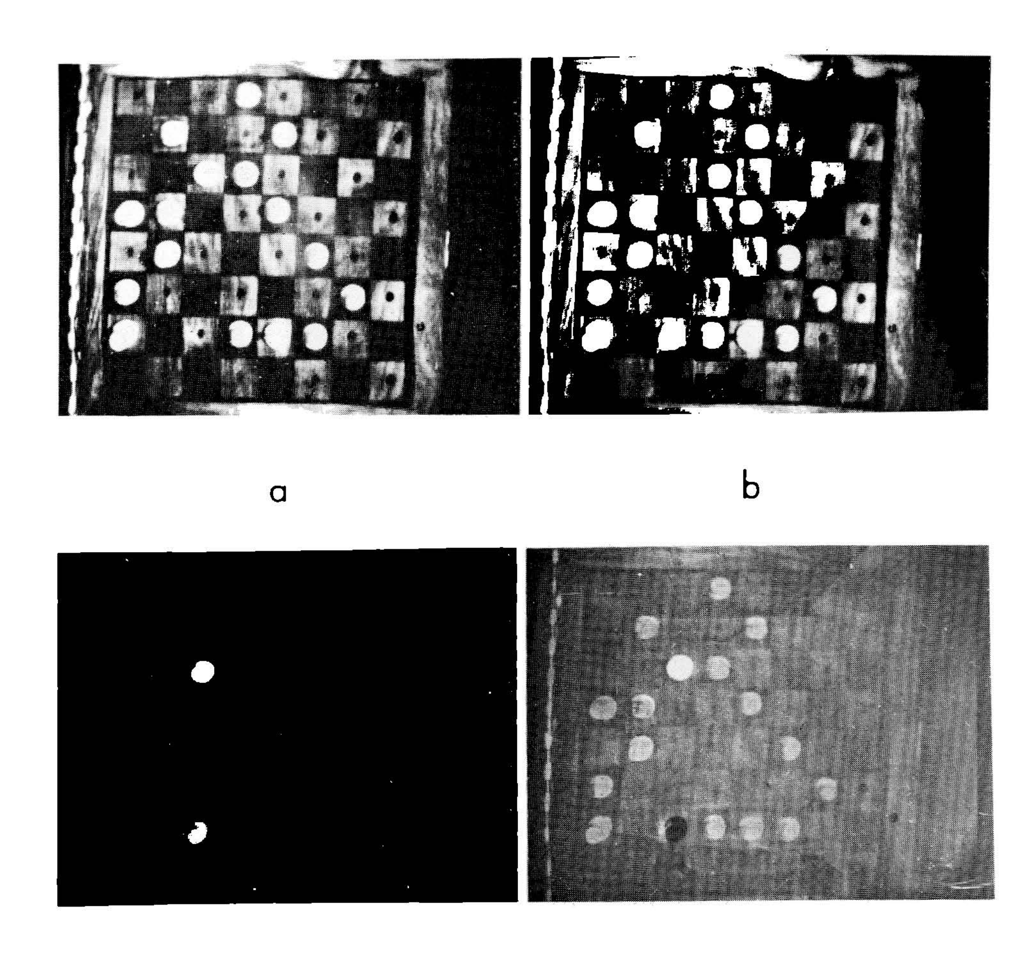


Figure 7 :
Real-time grey scale image subtraction.
Fig. 7a : image initially recorded on Phototitus with a given polarity.
Fig.7b does not result from an erasure but from a substraction procedure : the same picture being reprojected on the addressing side of the device with an opposite polarity bias applied.



c d

<u>Figure 8</u> : *Change detection*
Fig. 8a : original scene
Fig.8b : scene with an object displaced
Fig.8c : difference between the two scenes
The photographic plate being a quadratic detector both the removed and added objects appear in white ; the addition of a bias allows to remove, here, this ambiguity (Fig.8d); part of the initial image was preserved to show the background setting.

the gradient operation ; the latter yields a vector and one only displays its component along the direction of the shift. Omnidirectional second order differentiation can be displayed by subtracting for example from an initial scene recorded on Phototitus the same scene slightly blurred.

Figure 9b is the result of such an operation on a circular aperture (Figure 9a), the photographic plate being a quadratic detector only a white outline on a dark background is observed ; in figure 9c, an optical bias was added (§ II-A) to yield a grey background in order to show the two concentric white and black outlines that one would expect from a second order derivative.

Differentiation is a powerful image processing tool, indeed enhancement of contours and high spatial frequencies can thus be obtained ; for example, if instead of a complete differentiation, one still preserves part of the picture, small features will be enhanced in the resulting image as illustrated in figure 10.

c) Spectral image conversion

Phototitus can be considered not only as a display device but also as a pick-up device ; this feature is useful, in particular, for spectral image conversion : near ultra-violet or infra-red to visible conversion can readily be performed with, for example, Se compounds as the photoconductor. Another interesting possibility is X-ray to visible conversion ; indeed the photoconductors used

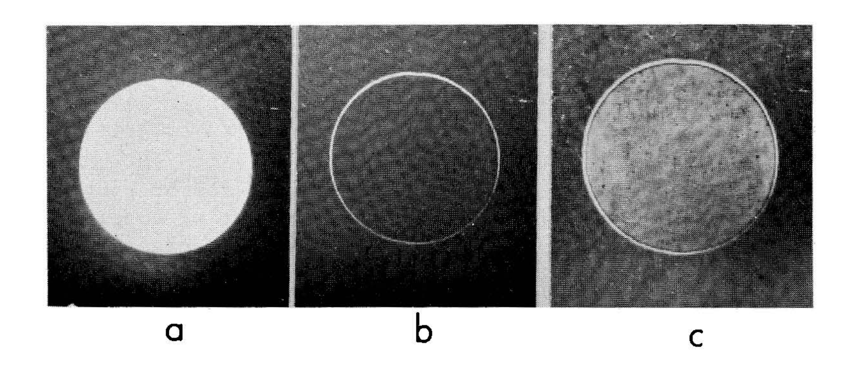


Figure 9 : *Second order omnidirectional differentiation.*
Fig. 9a : initial image (circular aperture).
Fig. 9b : differentiated image.
Fig. 9c : (enlarged) : differentiated image with an optical bias added to show the characteristic two concentric black and white outlines.

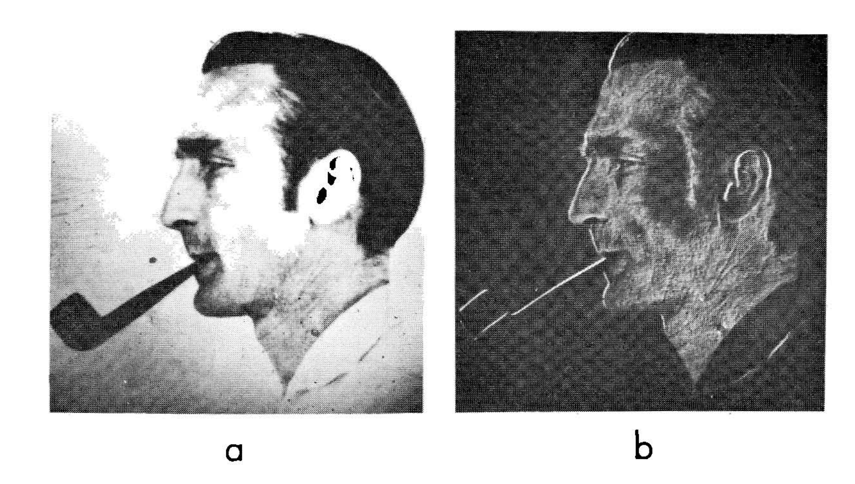


Figure 10 :
High spatial frequency enhancement by differentiation with part of the original image preserved.

(or usable) are photosensitive to X-ray, one X-ray photon yielding a great number of electron-hole pairs : for example a 25 KeV photon absorbed would generate about 100 electron-hole pairs [9] . Using a Phototitus originally designed to operate with visible addressing light, radioscopy feasibility experiments were carried out successfully and hard copies were simultaneously obtained quite simply.

The image conversion features of the device can be coupled with a real-time image processing in non-coherent or coherent light : positive to negative conversion, contrast enhancement, tomoscopy, change detection, etc.

Figure 11 shows an example of X-ray to visible conversion. An optical bias was used to obtain the negative (with some contouring) of the original radiogramme.

C) Coherent light applications

Coherent optical processing has been widely discussed in the literature , though actual industrial implementation has been long to come ; this was due in particular to the lack of suitable real-time incoherent to coherent convertors. Indeed, we can recall that, due to the speckle effect, data displayed on a scattering medium (e.g. paper) cannot, in general be directly processed in coherent light ; thus a sort of instant transparency is required. Since

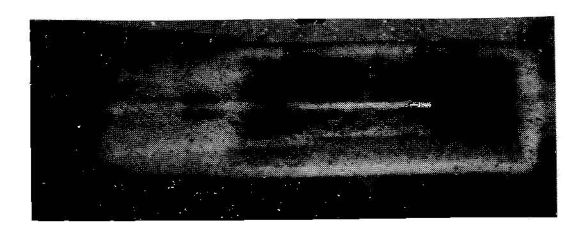


Figure 11 :
X-ray to visible conversion with Phototitus. An optical bias is used to obtain the negative (with some contouring) of the original radiogramme.

the DKDP target is a non scattering medium, Titus and Phototitus are light valves which can be used as real-time, high frame-rate, grey scale "incoherent to coherent convertors ". Furthermore, their storage capability is a valuable feature if scanned addressing is used (the whole processed image must be present simultaneously to perform a Fourier transform). Finally, the possibility of reducing the 0 order diffraction term can greatly increase the signal to noise ratio in a number of applications.

Of course several of the possibilities described in section III-B , such as differentiation, can be used in conjunction with coherent processing.

Real-time coherent optical processing allows a wide range of applications : a number of these applications were investigated and simulated to show that Titus and Phototitus were indeed well suited [4]. To quote a few : input for analog or digital optical multiplex memories, pattern recognition, complex filtering, speckle interferometry for increased resolution astronomical observation, non conventional imaging (reconstruction of acoustical holograms [6]), chirped Doppler radar, speach recognition, cine X-ray tomosynthesis using coded sources [10], etc.

These applications require for their implementation one or several of the characteristics of the devices enumerated above. To illustrate the suitability of Titus and Phototitus we will discuss real-time character recognition by auto-correlation which covers most of the problems encountered in coherent processing and whose requirements are quite demanding.

The processing is performed in an optical correlator similar to those described in the literature [11] with the device placed at the input plane. Photographic holographic filters of the various typewritten letters are first made using, for example, a

Titus tube in the input plane *). After development of the photographic plate, the latter is placed in the filter plane of the processor. In the recognition stage a typewritten text is illuminated with incoherent light, picked up by a TV camera and the resulting video signal fed into the Titus tube. Recognition is effected by auto-correlation in the "correlation plane".

S R V K S P A D F Z W N ? A D P A C / D

FILTRAGE SPATIAL



Figure 12 :
Automatic real-time character recognition using Titus.
Fig 12a : original text typed on paper.
Fig 12b : identification of the position of the three letters A present by the auto-correlation spots displayed, here, on a thresholded TV monitor.

Figure 12 shows an early result of character recognition using a Titus tube and the filter of the letter A. Figure 12 a shows a typewritten text ; in figure 12b, the position of the three letters A present are identified by the three auto-correlation spots displayed, here, on a thresholded TV monitor.

Using multiplex filters such as those described by Tréheux [12] reading speeds ten times faster than the fastest optical readers generally available should be possible.

---

*) The filters were made with the same as well as with a different tube from the one used in the recognition step. It can also be noted that when real-time implementation of the filters are not required photographic transparencies could be used.

IV. CONCLUSION

Judging by its possibilities, optical processing should, in principle, have a bright future next to electronic processing. In practice, however, it still suffers from a "credibility gap" due in particular to the scarcity of adequate devices required for implementing its potentialities.

It is hoped that the study and development of devices such as Titus and Phototitus will stimulate systems research in this field and help close this credibility gap and lead to a wide-spread use of optical processing.

ACKNOWLEDGEMENTS

Several of the applications described have originally been discussed with G. Marie and J. Donjon. The author is also grateful to F. Dumont for his help in some of the X-ray experiments and to D. Rossier for fruitful discussions. The author wishes to thank the Director of the Laboratoires d'Electronique et de Physique Appliquée for permission to publish this contribution.

REFERENCES

[1] G. MARIE Philips Res. Rep. 22 , 110 (1967)

[2] M. GRENOT, J. PERGRALE, J. DONJON and G. MARIE Appl. Phys. Lett. 21 , 83 (1972)

[3] G. MARIE, J. DONJON Proc. IEEE, 61, 942 (1973)

[4] G. MARIE, J. DONJON, J.P. HAZAN Pockels Effect Imaging Devices and their Applications in Advances in Image Pick up and Display Devices, Vol 1 225-301 (1974), B. Kazan ed., Academic Press New York.

[5] J. DONJON, F. DUMONT, M. GRENOT, J.P. HAZAN, G. MARIE J. PERGALE I.E.E.E. Trans. El. Dev. ED20 ,1037 (1973)

[6] J.P. HAZAN Proceedings of the 1974 Electro-Optics International Conference, Brighton, March 1974.

[7] D. ROSSIER, F. DUMONT Dependence of the Characteristics of a Photoconductor-Dielectric Optical Converter on Photogeneration and Charge Transfer Parameters of the Photoconductor, to be published in J. Appl. Phys., Jan. 1975.

[8] G. GROH, G. MARIE Optics Communication 2, 133 (1970)

[9] A.D. COPE, A. ROSE J. Appl. Phys. 25 , 240 (1954)

[10] H. WEISS Proc. of 74 International Optical Computing Conference Zurich , 9-11 April 1974.

[11] A. VANDER LUGT
I.E.E.E. Trans. Inf. Theory IT-10, 139 (1964)

[12] M. TREHEUX
Symp. Int. sur les Applications de l'Holographie
paper 13-3 (1970) (J.C. Viénot et al ed.), Univ. Besançon
Press.

COUPLING OF OPTICAL DEVICES AND WAVEGUIDES

H. Kogelnik

Bell Telephone Laboratories

Holmdel, N.J. 07733, U.S.A.

I. INTRODUCTION

At present a variety of devices and systems configurations are under consideration for fiber optical communications.[1] Early systems are likely to use multimode fibers with core diameters of about 50 μm to 100 μm. Later there may be applications for systems with large transmission bandwidths using single-mode fibers whose core diameters are typically 3 μm. The devices include detectors, modulators, light-emitting diodes (LEDs) which provide incoherent radiation, and lasers emitting coherent light. Some of these devices, such as junction lasers and modulators are becoming available in guided-wave form where the light is guided in thin planar films or strips. The interconnection or coupling of these various parts is the subject of this paper. We will try to illustrate the variety of coupling problems and describe approaches to solutions which have been found. Some of these solutions are quite ingenious, but we should not neglect to state that still simpler and more efficient coupling techniques are desired in almost all phases of optical communications and guided-wave optics.

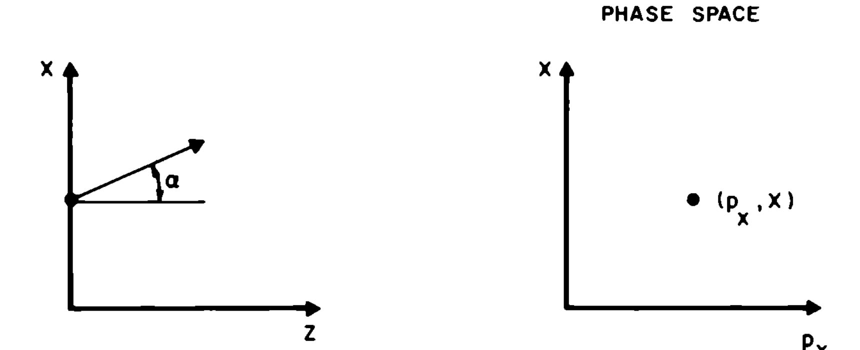
We will divide the discussion into two parts. The first deals with the coupling of multimode structures, where we are, essentially, concerned with incoherent light. The key example here is the coupling of an LED to a multimode fiber. The second part deals with the coupling of single mode structures where one is handling the coherent light from a laser. The coupling of a laser to a single mode fiber by a microlens is an example for this. Waveguide tapers, directional couplers, prism couplers and grating couplers are among the other techniques under exploration for the coupling of coherent light.

II. MULTIMODE COUPLING

A. Liouville's Theorem

The Liouville theorem is well known in statistical mechanics.

Its application to geometrical optics provides a powerful tool for the assessment of the coupling of multimode structures . [2,3,4] In geometrical optics the Liouville theorem deals with rays and with phase-space. In a given cross-section (z=const) of an optical system a ray is characterized by its position (x,y) and momentum (p_x,p_y) relative to the optic (z) axis. As indicated in Figure 1 the momentum $p_x(x,y,z)$ is related to the refrac-



$p = n \cdot \sin \alpha$

RAY POSITIONS: X, Y; RAY MOMENTA: p_x, p_y

Figure 1 :
Sketch of a ray in a cross-section z=0 of the optical system and its corresponding location in phase space

tive index n(x,y,z) and the ray angle α. $p_x = n.\ \sin\ \alpha$ (1)
and similarly for p_y. The phase space has four dimensions with the coordinate axes, x,y, p_x , and p_y. Its simplified 2D form is sketched in Figure 1. Each ray is represented by one point (x,y,$p_x$$p_y$) in phase space. This representation print moves around in phase space as the ray propagates through the optical system.

To describe a ray bundle we use the radiance (or brightness) $b(x,y,p_x,p_y,z)$ which is an appropriately normalized measure for the density of representation points (i.e., the density of rays of given location and momentum). The cross-sectional area and the cone angle of the bundle determine the volume V_p in phase space occupied by the rays. The total power P(z) carried by the bundle is given by an integral over V_p.

$$P = \int_{V_p} b(x,y,z,p_x,p_y)\ dx\ dy\ dp_x\ dp_y \qquad (2)$$

The first form of Liouville's theorem states [2] that the radiance is conserved from one cross-section to the next, i.e., that :

$$\frac{db}{dz} = 0 \tag{3}$$

where it is assumed that no rays are lost from the system. The second form of the theorem applies to the special cases where the radiance is uniformly distributed in V_p, i.e.

$$b(x,y,p_x,p_y) = B = \text{const} \tag{4}$$

For this condition, we obtain the result that V_p does not change with z,

$$V_p = \text{const} \tag{5}$$

which is the second form of Liouville's theorem. We should add that there exist several other forms for this theorem as well as several corollaries including Snell's refraction law and the Abbe sine-condition, but the above are sufficient for our present purposes. Note, also, that the shape of the occupied volume in phase space can change considerably from one cross-section to another. Figure 2 illustrates this for the 2D case of a ray bundle propagating in free space. One of the beauties of Liouville's theorem is its general validity. Within the limits of ray optics almost arbitrary refractive index distributions n(x,y,z) are allowed. Losses due to scattering, absorption or Fresnel reflection, however, are not taken care of and are assumed to be zero.

The assumption of constant B, Equation (4), provides a quite good description for the radiation characteristics of the light emitted by an LED. The power emitted by an LED of area A_{LED} into a solid acceptance angle Ω in free space (n=1) is, thus given by a simplification of Equation (2) in the form [1]

$$P = B \cdot A_{LED} \cdot \Omega \tag{6}$$

where B is usually measured in Watts/Steradian/cm^2.

B. Phase-Space Volume and Number of Modes

Let us, now, explore the relation between the volume V_p and the number of modes N supported by a structure. Consider, first a Gaussian laser beam [5] or "free-space mode" with a radius w_o at the beam waist and a far-field angle $\theta = \lambda / \pi w_o$. At the waist the beam area A is

$$A = \pi w_o^2 \tag{7}$$

and the solid far-field angle Ω is given by :

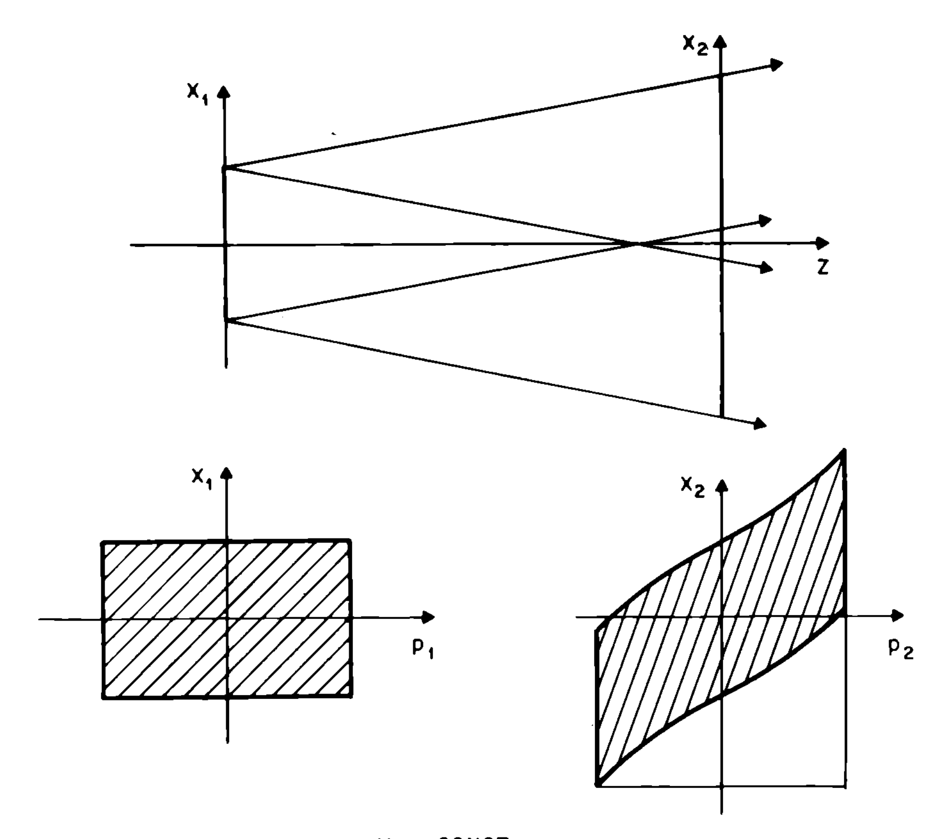


Figure 2 :
A ray bundle in free space and the evolution of its phase-space volume.

$$\Omega = \pi\theta^2 = \lambda^2 / \pi w_o^2 \tag{8}$$

The phase-space volume V_p of a gaussian beam is, therefore,

$$V_p = A \cdot \Omega = \lambda^2 \tag{9}$$

Next consider a multimode fiber with core radius a, core index n_1 and cladding index n_2 , as sketched in Figure 3. Total internal reflection at the core-cladding interface ceases when :

$$n_1 \cos\alpha = n_2 \tag{10}$$

This implies that the fiber will accept a maximum ray momentum of :

$$p_{max} = n_1 \sin \alpha = \sqrt{n_1^2 - n_2^2} \qquad (11)$$

This is also known as the "numerical aperture" of the fiber.[1] The corresponding phase-space volume is :

$$V_p = A.\Omega = \pi^2 a^2 p^2_{max} = \pi^2 a^2 (n_1^2 - n_2^2) \qquad (12)$$

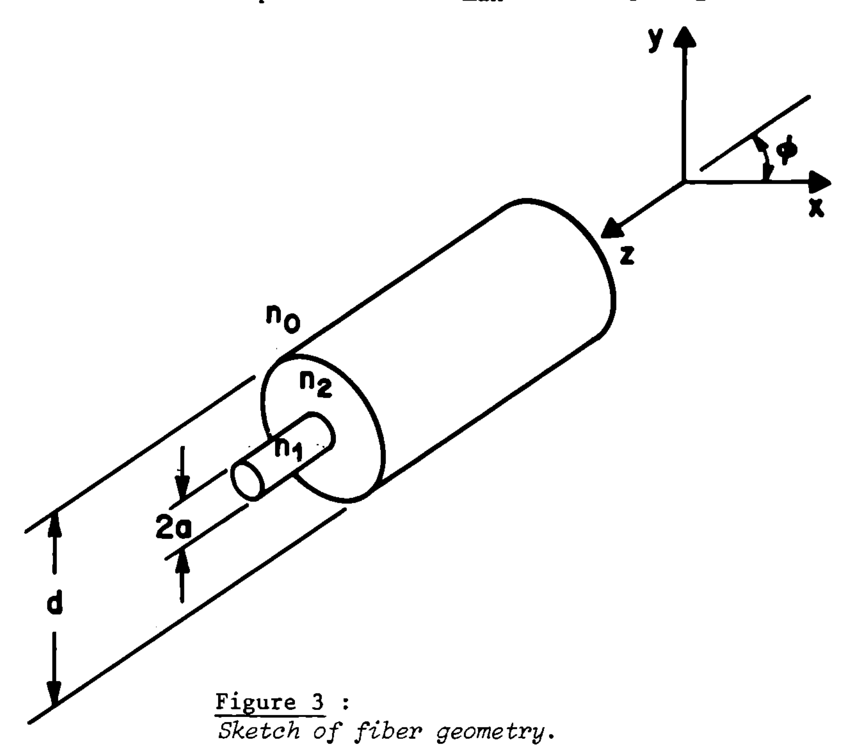


Figure 3 :
Sketch of fiber geometry.

As the number N of guided modes allowed in the fiber is given by [1]

$$N = 2\pi^2 (n_1^2 - n_2^2) a^2/\lambda^2 \qquad (13)$$

we have the simple relation :

$$\lambda^2 N = 2V_p \qquad (14)$$

Considering that there are two polarization states we note, again, that each propagating mode occupies the volume of λ^2 steradian.cm^2 in phase-space. This establishes a connection to wave-theory. The conservation of V_p in ray optics becomes the conservation of the number of modes in the language of wave-optics.

C. Coupling of an LED to a Multimode Fiber.

The usual multimode fiber accepts only a fraction of the light

emitted by an LED. This is because of the relatively small acceptance angle of the fiber, which, in the typical case of a = 50μ m and a relative index difference of 10%, accepts a phase-space volume V_p of about 2×10^{-5} sr.cm^2 . The maximum power that can be coupled from the LED to the fiber follows from Equations (6) and (12) as :

$$P = B \cdot V_p = \pi^2 B a^2 (n_1^2 - n_2^2) \qquad (15)$$

This upper limit is a direct consequence of Liouville's theorem and the fact that the phase-space volume V_p of the fiber is smaller than that of the LED. It is independent of the optics between the LED and the fiber. The only LED parameter appearing in this formula is the LED radiance B.

To increase P one can, in principle, increase the index difference $(n_1 - n_2)$. But systems requirements [1] call for a compromise here, because the group-delay spread in the fiber also increases with $(n_1 - n_2)$.

To couple power levels of the order of 2 mWatts into a multimode fiber with an index difference of 10 percent, one needs small area (≲ 50 μm diameter) LEDs capable of high current-densities (≳ 10^4 A/cm^2) that can provide a high radiance of about 100 W/sr/cm^2 . Burrus et al [6,7] have reported LEDs with such characteristics. They have employed a GaAlAs double-heterostructure grown by liquid-phase epitaxy, which is the same material system as that used for cw room-temperature injection lasers. [8,9] One of the advantages of this LED is that its emission peak can be adjusted to any wavelength in the range from 0.75 - 0.9 μm by choosing the proper Al concentration in the light emitting layer. This is a wavelength range where the losses of optical fibers tend to be low.[1] Figure 4 shows a sketch of the cross-section of such an LED. A small light emitting area is controlled by the small contact area shown in solid black. A drive current of about 150mA was needed to produce a radiance of 100 W/sr/cm^2 .

Figure 4 also shows the coupling technique due to Burrus, where a well is etched into the GaAs substrate. The fiber end is brought into this well into the close proximity of the primary light emitting area, and is secured by an epoxy resin. Another advantage of the GaAlAs double-heterostructure LED should be mentioned here, and that has to do with the absorption of the n-type AlGaAs layer situated between the fiber and the light emitting area. The Al concentration in that layer can be adjusted for low absorption losses.

The Burrus technique is a simple and direct coupling scheme. It is natural to ask the question whether improvements over this can be gained by more complicated schemes employing lenses or waveguide tapers between the LED and the fiber. As long as the radiance B is constant it is clear from Liouville's theorem that nothing can be gained by imaging a larger area LED onto the fiber. On the other hand, one can consider the coupling of a smaller area LED to the fiber by means of a lens or conical taper. As long as Equation (15) holds, there is no decrease in the maximum

power if B is constant and if the optics is used to match the V_p of the fiber and LED. Such a scheme would imply a better coupling efficiency and reduced LED drive power, but it would also make the coupling elements more complicated and more costly.

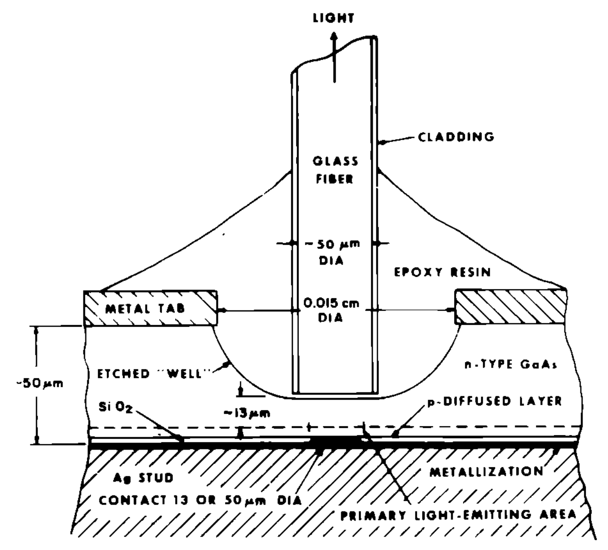


Figure 4 :
Cross-Section of AlGaAs LED coupled to an optical fiber. (From Reference 6)

III. SINGLE MODE COUPLING

For the purposes of this discussion we are interested in four different optical structures that can support single modes. These are light beams [5] propagating in free space, dielectric film waveguides [10,2], planar dielectric strip waveguides [2,11,12] and fibers. Techniques which allow the simple and efficient coupling between these structures are still evolving, but a considerable variety of coupling devices has already been proposed. Table 1 presents a selection of the methods studied for coupling the various structures and waveguides. For example the usual laboratory technique to couple a laser beam to a fiber is to focus that beam directly onto the fiber core with a lens. An alternative which has been studied for the same purpose (and which is not listed in the table) is to use two steps. [13] In the first step, the beam is fed into a guiding film via a grating coupler, and in the second step a planar-film lens is used to focus the light onto the fiber. From the variety of proposed solutions we have singled out some of the more prominent ones for the following discussion. These are the use of microlenses, of waveguide tapers, of directional couplers, and of prism and grating couplers.

	BEAM	FILM	STRIP	FIBER
FIBER	DIRECT FOCUSING		MICRO LENSES	SPLICE
STRIP	2 STEPS?	PLANAR PRISM OR GRATING	TAPERS DIRECTIONAL COUPLERS	
FILM	PRISM GRATING	OVERLAPPED TAPERS DIRECTIONAL COUPLERS		
BEAM	MODE MATCHING			

Table I

A. Microlenses

In the coupling of lasers to fibers there are no fundamental limitations such as the Liouville theorem as the laser output is usually available in a single transverse mode. The practical difficulties encountered here depend very much on the particular laser structure used and on the core radius and the numerical aperture of the fiber. Figure 5 shows a sketch of a typical AlGaAs

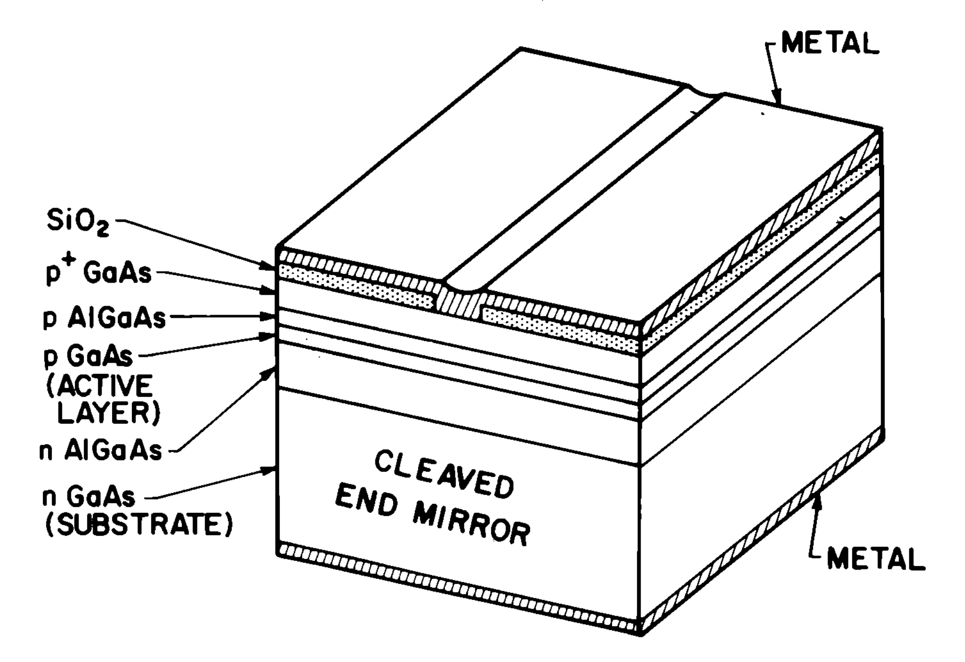


Figure 5 :
Sketch of stripe-geometry AlGaAs double-hetero-structure injection laser (from Reference I)

double heterostructure (DH) injection laser, where the lateral confinement is obtained by an oxide-masked stripe. Other techniques used to produce a stripe geometry include proton-implantation or the growing of mesa structures. Present day low-threshold cw DH lasers [8,9] produce rectangular output beams with cross-sections of about 0.6 μm×4 μm at the cleaved end mirror.

When the divergence angle of the laser is less than the acceptance angle of the fiber one can obtain simple and direct coupling by aligning the fiber core with the laser stripe [14,15]. By optimizing the air-gap spacing between laser and fiber, efficiencies of about 8% have been achieved for the coupling of a DH injection laser with a divergence angle of α_y 30° to a 4 μm core single mode fiber. [15]

To improve the coupling of lasers with large divergence angles Cohen and Schneider [16] have proposed the use of microlenses. As indicated in Figure 6 these lenses can be attached either to the

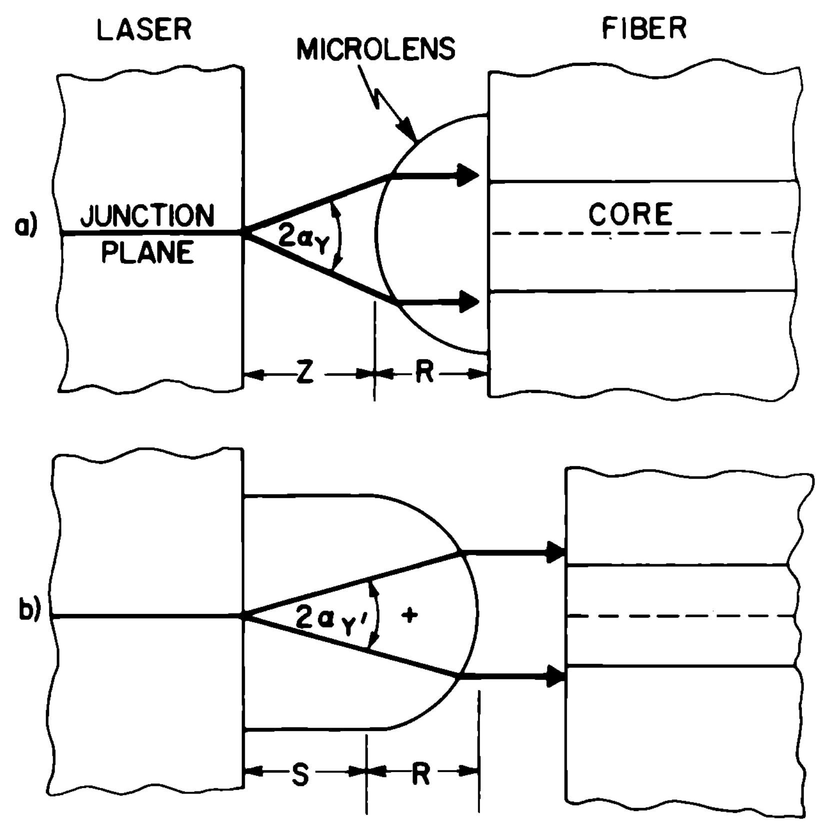


Figure 6 :
Schematic of laser-fiber coupling via microlenses. (From Reference [16])

fiber or the core, and they help to collimate the light from the laser. A simple ray analysis [16] shows that lenses of a few μm

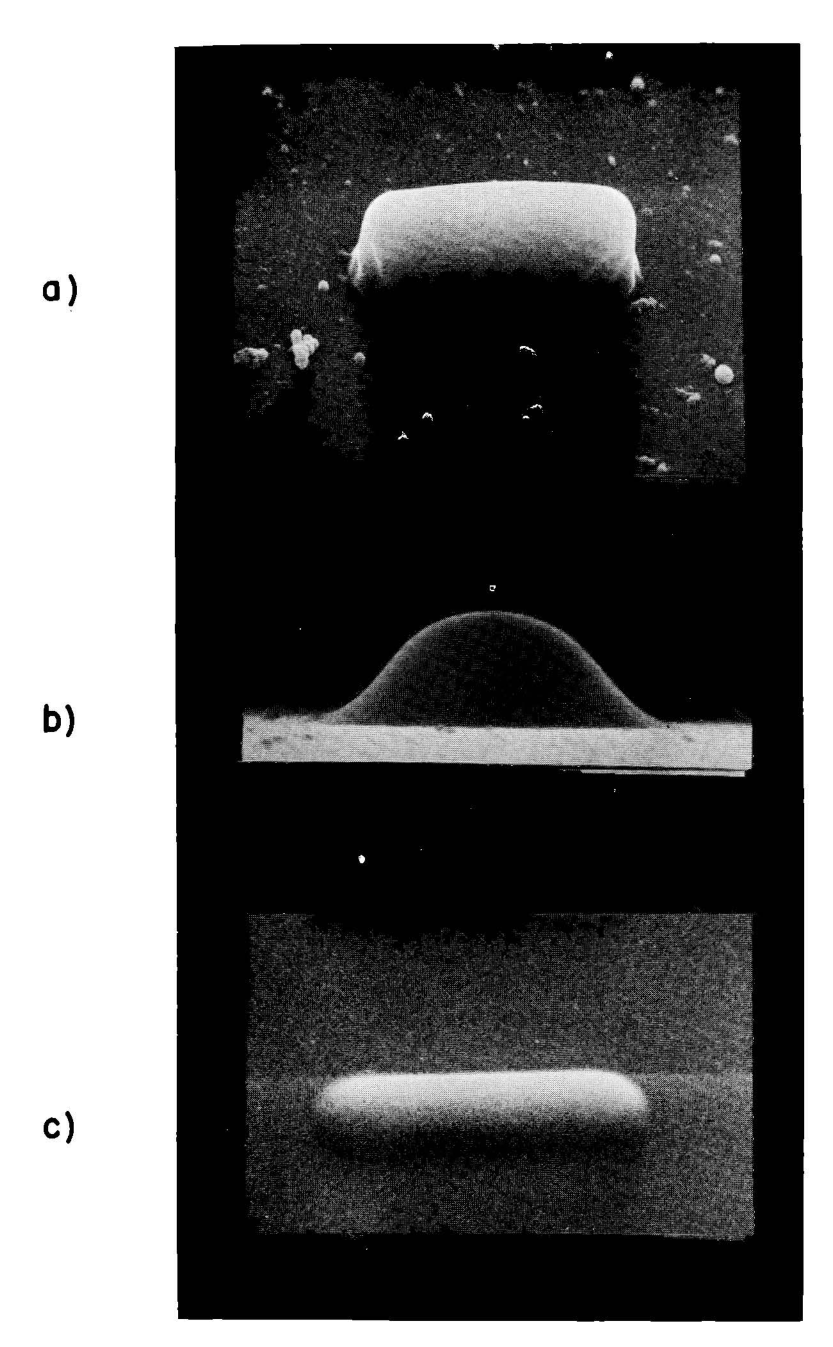


Figure 7 :
Scanning electron micrograph of cylindrical microlenses. (From Reference I6)

in size and diameter are required. The use of mode matching formulas [5] yields a similar answer. Cohen and Schneider have made such microlenses from photoresist which was deposited and exposed directly on the fiber end. Because of the rectangular laser output cylindrical lenses were found to be more effective than spherical lenses. For lasers and single-mode fibers of the same structure as above a coupling efficiency of 23% was obtained with spherical lenses and 34% with cylindrical lenses. Figure 7 shows a scanning electron micrograph of the cylindrical photoresist lenses which had a length of about 25 μm and a diameter of 7 μm.

Kato [17] has suggested that spherical microlenses be formed by thermal melting of the ends of multimode fibers. The heating can be done with a hydrogen-oxygen torch, for example, and surface tension will help to create a spherical surface and reduce the surface roughness at the fiber end. Brackett [18] has recently used this technique to fabricate lenses with radii between 65 μm and 80 μm on multimodes fibers with a core diameter of 80 μm. In coupling these to a DH injection laser with half angles of 45°× 6° he has achieved coupling efficiencies of 62%. Figure 8

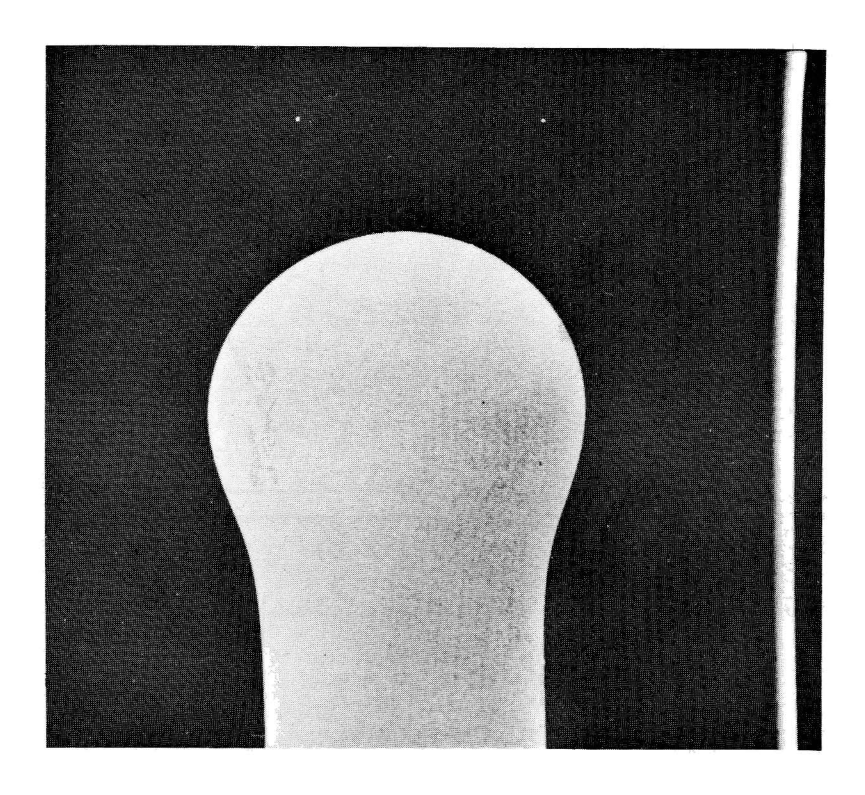


Figure 8 :
Scanning electron micrograph of cylindrical microlenses.(From Reference I6)

shows a scanning electron micrograph of such a fused microlens.

Short sections of fibers with a suitably designed parabolic refractive index profile (SELFOC) can also serve as microlenses . Kobayashi et al [19] have recently reported the use of cylindrical and spherical lenses of this type to couple junction lasers to SELFOC fibers.

The coupling problem due to the large divergence angles of DH injection lasers can also be attacked at its root. In this connection, a solution has been pointed out by the recent proposals of DH lasers with a thick guiding structure for optical confinement [20-24] .

B. Tapered Waveguides

Single-mode dielectric waveguides such as strip guides or fibers can be connected by tapered sections which provide a transition from one waveguide cross-section to the other. In principle, we know that all tapered dielectric guides radiate and are, therefore, lossy. But these losses can be kept sufficiently low if the tapers are made gradual enough.[2] This usually requires taper lengths much larger than a cross-sectional dimension.

Strip guides of different width can be connected by a strip guide taper of gradually changing width. Ostrowsky et al [12] have made a similar transition tapering from 25 μm to 5 μm in width in order to connect a film guide to a strip guide. This was used as part of a two-step coupler (prism coupler, taper) to feed a laser beam into a strip guide.

Tien et al [25] have coupled two different film guides by a bridging film with ends tapered in thickness. Figure 9 show this

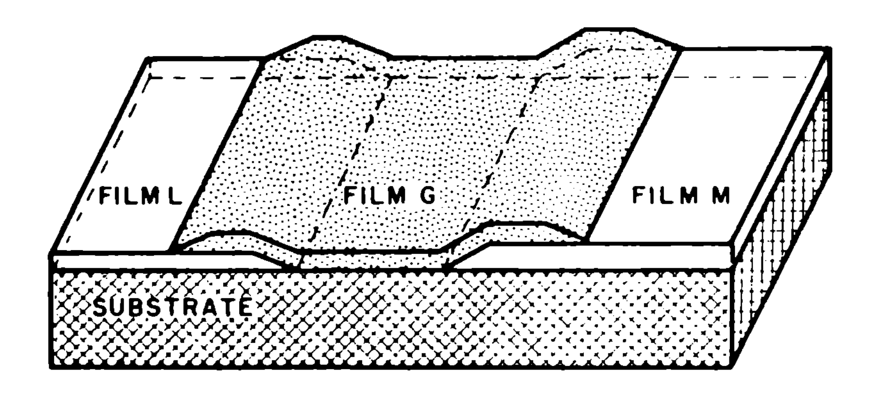


Figure 9 :
Sketch of tapered film-bridging interconnection. (From Reference 24).

interconnection. The tapered film ends required for this form naturally during the fabrication process.

A tapered film end can also be used to couple light from a film guide to the substrate [26] . Figure 10 shows how a zig-zag angle

of a guided mode changes due to the taper and light is coupled to the substrate by total internal reflection. By reciprocity light can be coupled into the film via this taper.

C. Directional Couplers

A directional coupler is a device familar in microwave technology. An optical analog of this has been proposed for the coupling

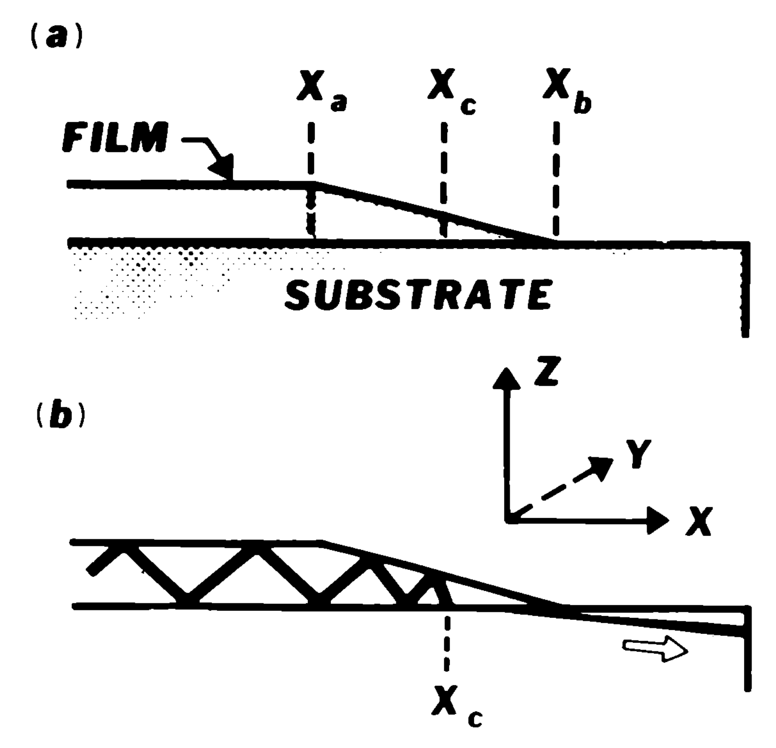


Figure 10 :
Sketch of ray paths in tapered film coupler. (From Reference 25)

of the guided-wave devices of integrated optics, and also as a 3dB coupler or power divider for more complicated optical circuitry [11 12]. Figure 11 shows a sketch of an optical directional coupler made from two (single-mode) strip guides (with a typical width of a few μm) on a glass substrate. The guides approach each other and can exchange energy via the evanescent fields of their modes. The coupling between the two guides is strongly (exponentially) dependent on the guide separation, which is typically 1 μm. Complete transfer of power (e.g., from P_i to P_{02}) is possible, and requires interaction length of about 100 μm. The proper functioning of this device requires synchronism, i.e., the matching of the propagation constants of the two waveguides. The fabrication of these devices usually requires the use of scanning electron beam writing techniques for good control of edge roughness

($\approx$500 Å) and of the waveguide dimensions. The first experimental fabrication of directional couplers has been reported by Goell [11] and Ostrowsky et al. [12]

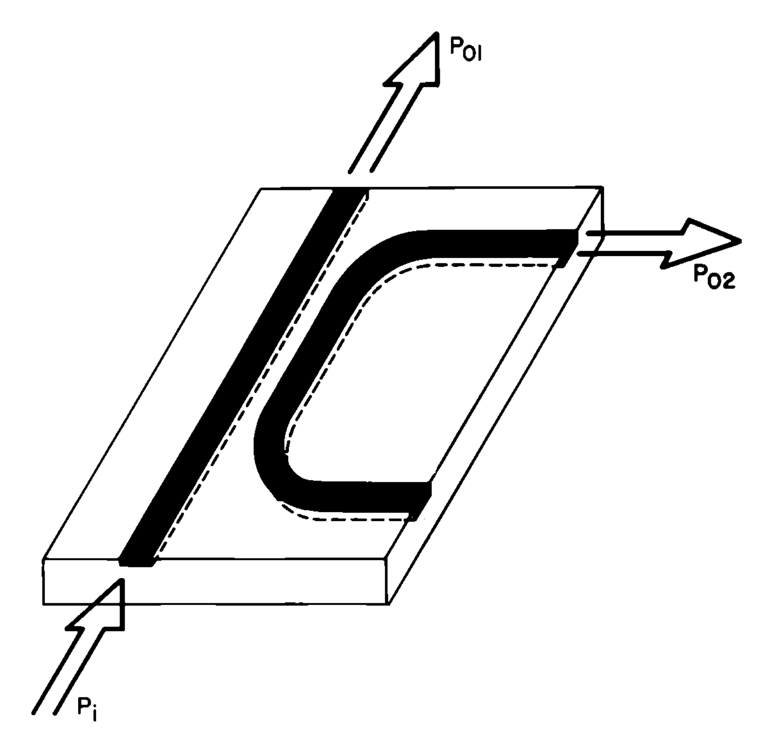
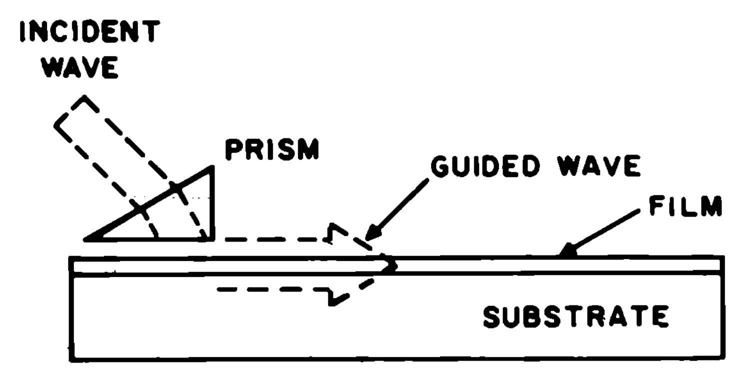


Figure 11 : *Directional coupler formed by dielectric strip guides.*

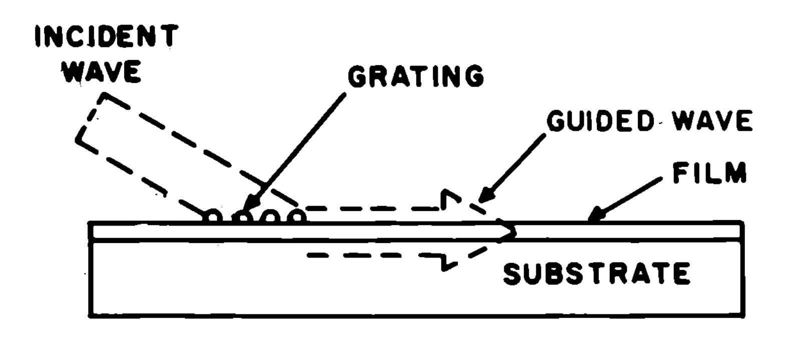
D. Prism and Grating Couplers

Prism and grating couplers are used to feed the light from a (free-space) laser beam into a planar film waveguide. Figure 12 shows these two devices. To form a prism coupler [10,27] one uses a prism of a refractive index n_p that is larger than the index n_f of the film ($n_p > n_f$). As sketched in Figure 12a this prism is brought close to the film with a separation (or "coupling gap") of the order of 1000 Angstroms. The laser beam is incident through the prism and coupling takes place via the evanescent fields in the gap.

To make a grating coupler [28, 29] one machines a grating on the film surface, as indicated in Figure 12b. Gratings with a few 1000 lines/mm are required, and are usually made by (holographic) laser interference techniques. The grating coupler works by diffracting the incident beam into the desired waveguide mode.



(a) PRISM COUPLER



(b) GRATING COUPLER

Figure 12 :
Schematic for prism and grating couplers.

The overall characteristics of the prism and grating couplers are quite similar. For uniform coupling gaps or uniform grating strength the maximum predicted coupling efficiencies are about 80% and to achieve high efficiencies one has to match the width w_B of the incident beam to the leakage length ℓ of the coupler ($w_B \approx \ell$). The latter is a measure for the exponential decay (exp $(-z/\ell)$) of a guided film mode in the coupler in the absence of an incident beam.

Another requirement for efficient coupling is, again, the phase-matching or synchronism of the incident beam and the guided mode it couples to. We achieve synchronism by choosing the correct angle of incidence θ with respect to the film normal. We will illustrate this by means of the wave-vector diagrams in the following figures, where we plot the propagation constant β of the guided wave and the wave vectors $k_s = 2\pi n_s/\lambda$ and $k_p = 2\pi n_p/\lambda$

of the plane waves in the substrate or prism reepectively.
Figure 13 shows the wave-vector diagram for the prism coupler.

PRISM COUPLER

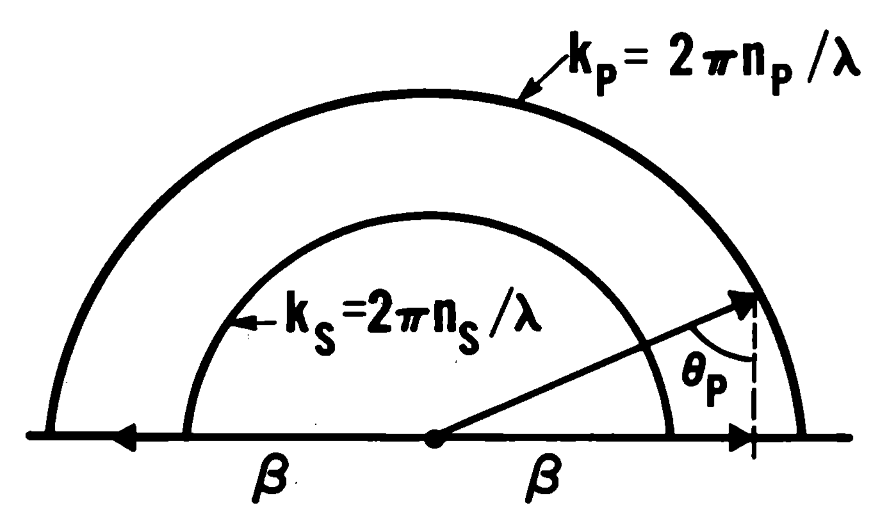


Figure 13 :
Wave-vector diagram for prism coupler.

The β of the guided modes is larger than k_s. The smaller circle is the locus of all allowed wave-vectors in the substrate. Synchronism requires that the in-film component (parallel to β) of a k-vector be equal to β, and this is clearly not achievable with the substrate waves. The larger circle represents the waves in the prism. Here synchronism is achieved for an injection angle θ_p in the prism which is sketched in the figure and is given by :

$$\beta = k_p \sin \theta_p \tag{16}$$

The wave-vector diagram for a grating coupler is shown Figure 14. Here we associate a wave-vector

$$k = 2\pi/\Lambda \tag{17}$$

with the grating where Λ is the grating period. Synchronism is achieved when :

$$\beta = K + k_s \sin \theta_s \tag{18}$$

for a wave incident at an angle θ_s in the substrate. The figure shows the case of forward coupling where the in-film component of k_s points in the same direction as β. The generation of unwanted grating orders can reduce the efficiency of this coupler. This

GRATING COUPLER

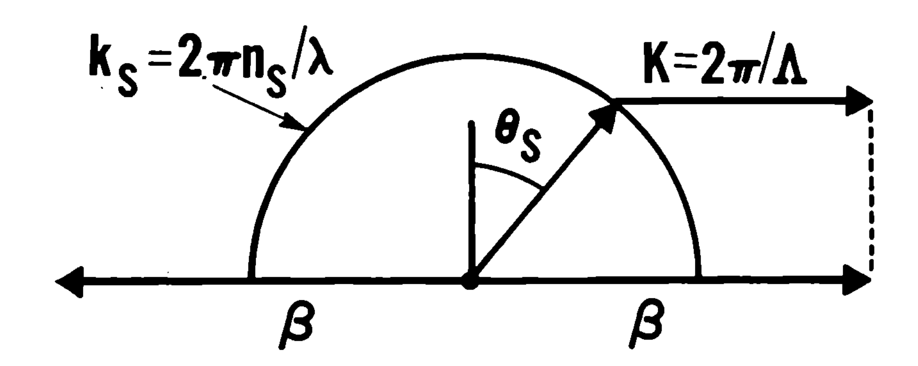


Figure 14 :
Wave-vector diagram for forward grating coupler.

problem is overcome in the reverse coupler [29] where β and the in-film k_s component point in different directions. This is illustrated in Figure 15. The reverse coupler requires larger K-vectors

REVERSE COUPLER

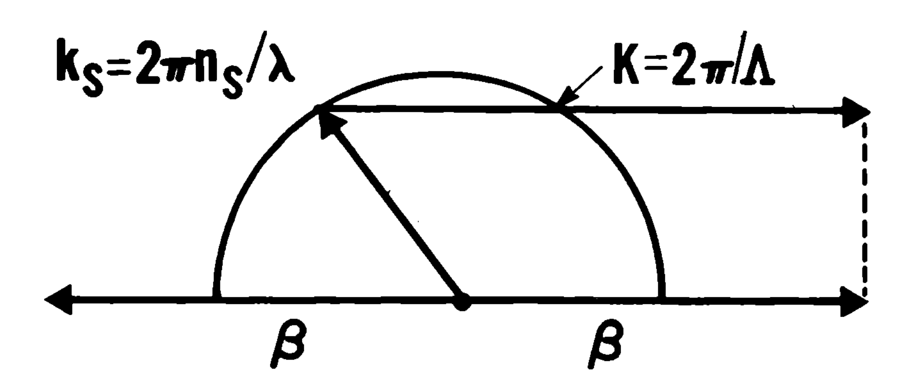


Figure 15 :
Wave-vector diagram for reverse grating coupler.

and, therefore, more grating lines per mm (up to about 5000), but the unwanted higher grating orders are generated beyond cut-off. Dalgoutte [30] has recently fabricated such couplers which had efficiencies exceeding 70%.

The prism and grating couplers can also be used in planar form to couple light from film guides to planar strip guides. The geometry of these structures is essentially the same as that shown in Figure 12 if we regard these as sketches of a top view perpendicular to the film plane.

IV. CONCLUSION

While the techniques for the coupling of optical devices and waveguides are still evolving there is already available a considerable variety of coupling methods. In experimental tests some of these have provided coupling with efficiencies as high as 60-70 percent. Some of the aims of present work are to evolve coupling techniques which are flexible, have improved efficiency and greater simplicity.

REFERENCES

[1] S.E. Miller, E.A.J. Marcatili and T. Li, Proc. I.E.E.E. 61, 1703 , December 1973.

[2] D. Marcuse, *Light Transmission Optics*, Van Nostrand Reinhold, Co., New York 1972.

[3] D. Marcuse, Appl. Optics, 10, 494, March 1971

[4] M.C. Hudson, Appl. Optics, 13 , 1029, May 1974

[5] H. Kogelnik and T.Li, Appl. Optics, 5 , 1550, October 1966

[6] C.A. Burrus and B.I. Miller, Optics. Comm., 4, 307, December 1971.

[7] C.A. Burrus, Proc. I.E.E.E., 60 , 231, February 1972.

[8] Zh. I. Alferov, V.M. Andreev, E.L. Portnoi and M.K. Trukan Sov. Phys. Semiconductors, 3, 1107, March 1970.

[9] M.B. Panish, I. Hayashi and S. Sumski, Appl. Phys. Lett., 16 , 326, April 1970.

[10] P.K. Tien, R. Ulrich and R.J. Martin, Appl. Phys. Lett., 14, 291, May 1969.

[11] J.E. Goell, Appl. Optics, 12 , 729, April 1973.

[12] D.B. Ostrowsky, M. Papuchon, A.M. Roy and J. Trotel, Appl. Optics, 13 , 636 , March 1974.

[13] L.P. Bovin, Appl. Optics, 13 , 391, February 1974.

[14] D. Schicketanz and J. Schubert, Optics Comm. 5 , 291, July 1972.

[15] L.G. Cohen, BSTJ, 51, 573 , March 1972.

[16] L.G. Cohen and M.V. Schneider, Appl. Optics, 13 , 89, January 1974.

[17] D. Kato, J. Appl. Phys., 44,2756,June 1973

[18] C.A. Brackett, J. Appl. Phys. , to be published.

[19] K. Kobayashi et al, Proc. Nat. Convention of Inst. El . Eng. Commn. of Japan, July 1974.

[20] H. Kressel, H.F. Lockwood and F.Z. Hawrylo, J. Appl. Phys., 43 , 561,February 1972.

[21] G.H.B. Thompson and P.A. Kirby, I.E.E.E. , J. Quantum Elect., QE-9 , 311,February 1973.

[22] T.L. Paoli, B.W. Hakki and B. Miller , J. Appl. Phys. 44 , 1276 , March 1973.

[23] M.B. Panish, H.C. Casey , Jr., S. Sumski and P.W. Foy, Appl. Phys. Lett., 22 , 590 , June 1973.

[24] G.H.B. Thompson and P.A. Kirby, Elec. Less., 9,295, June 1973.

[25] P.K. Tien, R.J. Martin anf G. Smolinsky, Appl. Optics, 12 , 1909 , August 1973.

[26] P.K. Tien and R.J. Martin , Appl. Phys. Lett., 18 , 398, May 1971.

[27] J.H. Harris and R. Shubert, Conf. Abstracts , URSI Spring Meeting , April 1969, Page 71.

[28] M.L. Dakss, L. Kuhn, P.F. Heidrich and B.A. Scott. Appl. Phys. Lett. , 16 , 523 June 1970.

[29] H. Kogelnik and T.P. Sosnowski, BSTJ, 49, 1602,September 1970

[30] D. Dalgoutte, Optics. Comm., 8 , 124, June 1973.

LIGHT MODULATORS AND MODULATOR MATERIALS

I. P. Kaminow
Bell Laboratories Crawford Hill Laboratories
Holmdel, N.J. 07733 U.S.A.

I. INTRODUCTION

Various methods have been proposed for modulating light (Fig 1).

MODULATION METHODS

(1) MECHANICAL (∿10 kHz)

(2) DIRECT MODULATION OF SOURCE :
SEMICONDUCTOR LASER OR LED (∿300 MHz)

(3) ELECTROOPTIC EFFECT ($\Delta n \sim E$)

(4) ACOUSTOOPTIC EFFECT ($\Delta n \sim S$)

(5) MAGNETOOPTIC EFFECT ($\Delta n \sim M$)

(6) OPTICAL WAVEGUIDE DEVICES

Figure 1 :

Mechanical methods, such as moving mirrors or shutters, are limited by inertia to frequencies less than 10 kHz. Laser and light emitting diodes can be current modulated at high frequencies, but other lasers require external modulators based on various microscopic effects. The operation of these electro-, acousto- , and magnetooptic devices is reviewed elsewhere. [1,2,3] In order to improve the performance and ease of fabrication of these bulk modulators, it is profitable to employ optical waveguide techniques. Much of this talk is based on a recent review of such devices.[4]

II. BULK DEVICES

The operation of an electrooptic modulator is illustrated in Fig 2. The modulating field V/d produces a proportional index

ELECTROOPTIC MODULATOR

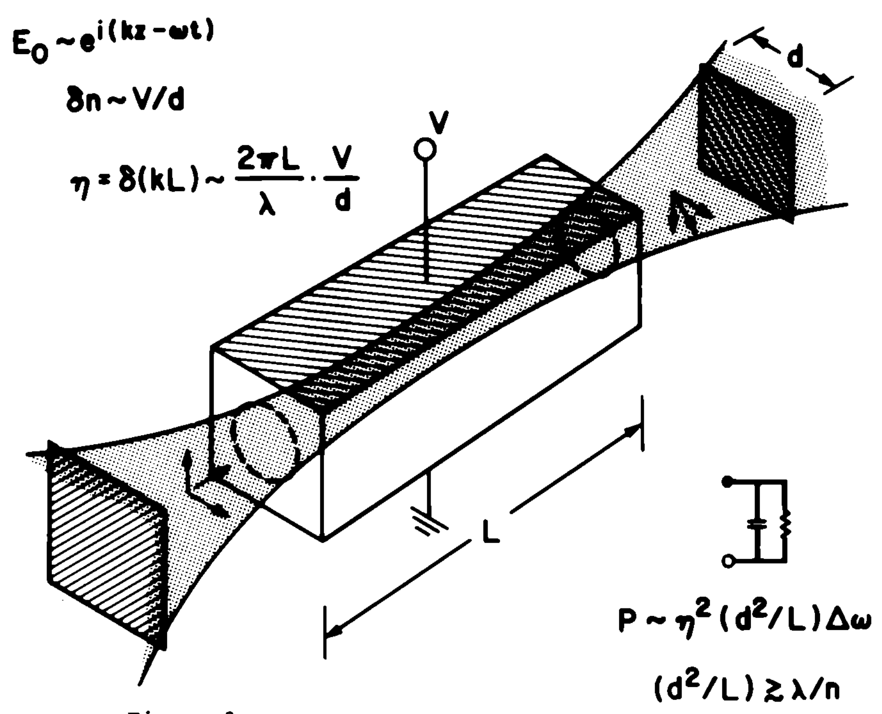


Figure 2 :
Bulk Electrooptic Modulator.

change δn. The modulator has capacitance C and is normally placed in parallel with resistance R(50Ω) to achieve bandwidth $\Delta\omega=(\pi RC)^{-1}$. Then the modulating power P is proportional to $\Delta\omega$, to the peak phase shift squared,η^2, and to a geometrical factor (d^2/L) which is limited by diffraction to a value greater than $4\lambda/\pi n$. The waveguide devices to be described remove the diffraction limit and allow very small values of $(P/\eta^2\Delta\omega)$.

Another mean for obtaining a broadband circuit is shown in Fig. 3. [5] The $LiTaO_3$ crystal is included in a matched (50Ω) microstrip transmission line. The only limit on modulating frequency ($\sim$2 GHz) is imposed by the difference in transit times of light and modulating waves through the crystal. [1,5] The modulating power is about 1mW/MHz of bandwidth for 100% modulation at $\lambda \gtrsim 0.5$ μm.

A recently developed $LiNbO_3$ modulator [6] operating at 1.06μm required 10 mW/MHz for 100% modulation because $P \sim \lambda^3$.

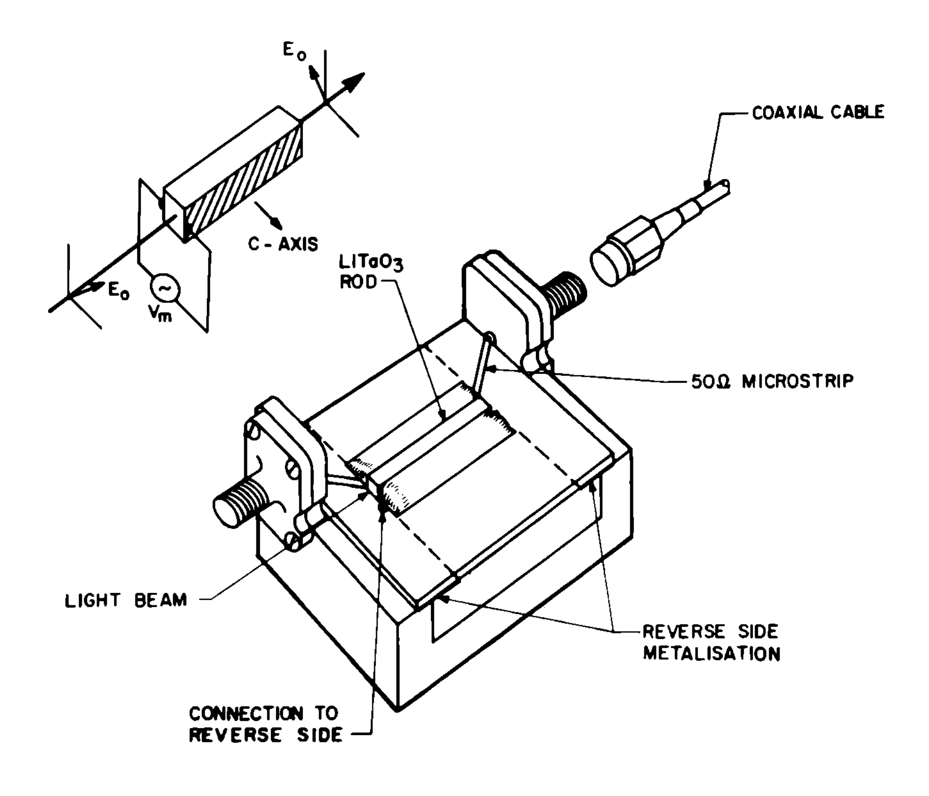


Figure 3 :
Microstrip Modulator .[5]

An acoustic wave in a solid deflects an optical beam through an angle $2\theta_B$ as in Fig. 4, where :

$$\theta_B \approx \frac{\Omega c}{2n\omega V} ,$$

with Ω the acoustic frequency, c the velocity of light, n the refractive index, ω the radian optical frequency, and V the acoustic velocity. Turning the acoustic beam on or off, turns the diffracted beam on or off. Varying Ω scans the diffracted beam. In either case, Ω must be large enough to separate the beams : $\theta_B > \lambda/D$, the angular beam width, where D is the beam diameter.

A bulk As_2Se_3 acoustooptic modulator designed to operate at 1.06 μm with an optical fiber output required 1mW/MHz for 70% modulation. [7]

The magnetooptic or Faraday effect produces a rotation ϕ of the polarization, proportional to the component of magnetization along the direction of propagation. (Fig 5). The Faraday effect is large near the Fe^{3+} absorption edge at ∿ 1μm in YIG and a high speed modulator has been demonstrated by varying $\underline{M}.\underline{k}$ using an external coil. [3]

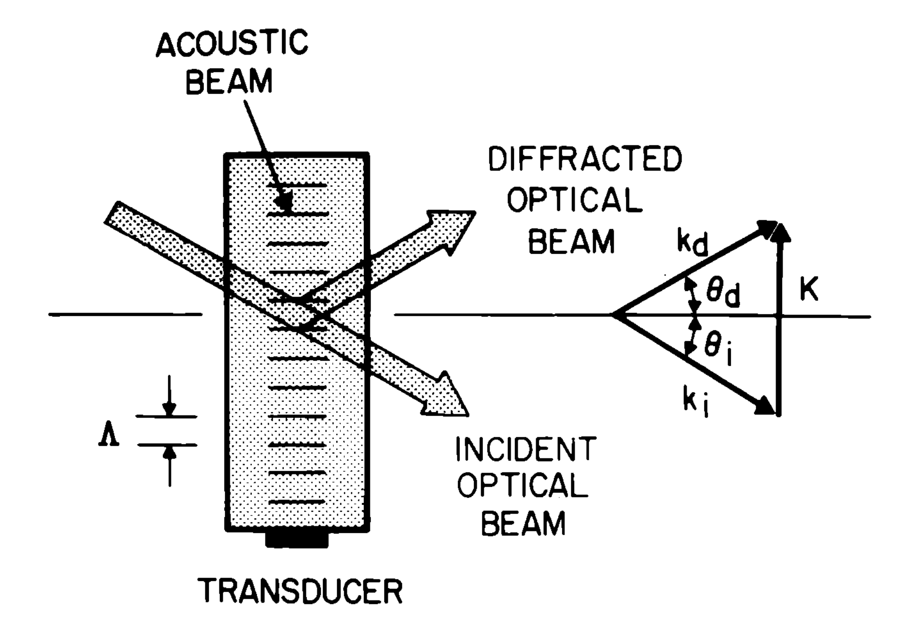


Figure 4 :
Bragg Deflection by the Acoustooptic Effect.

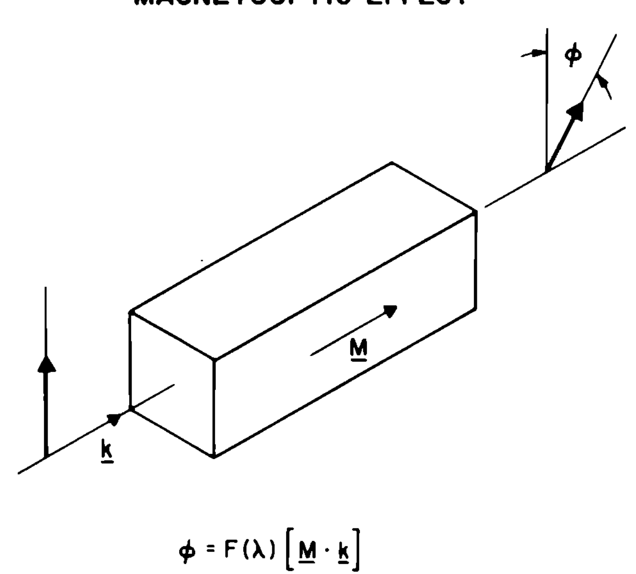


Figure 5 :
Magnetooptic Effect.

III. OPTICAL WAVEGUIDE DEVICES

An optical waveguide, based on total internal reflection, can be made by sandwiching a film of high index, n_1, between a substrate, n_2, and superstrate, n_3, of lower index, where :

$$n_1 > n_2 \geq n_3$$

as illustrated in Fig. 6. The zig-zag ray may be polarized normal to (transverse electric, TE, mode) or in the plane of (transverse magnetic, TM, mode) the figure.

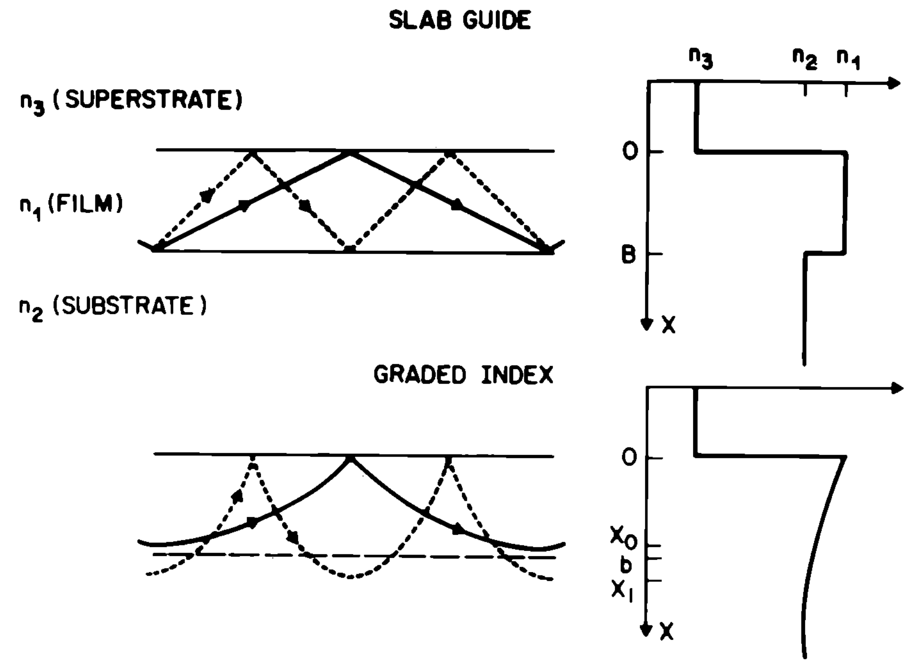


Figure 6 :
Slab and Graded Index Waveguides.

Various techniques have been developed for achieving waveguiding films in electro-, acousto- and magnetooptic materials. Some methods, such as epitaxial growth, yield slab guides consisting of a uniform film and others, such as diffusion, yield a graded index layer. For the slab guide, the rays of both low and high order modes turn at the boundaries x = 0,B. For the graded index guide, the turning points, $x_{0,1}$, are deeper for higher order modes.

Guided light can neither leave nor enter through the film-substrate or film-superstrate boundaries. The mode may be excited by injecting light from the end or by perturbing a boundary : with air as the superstrate, a periodic grating may be etched on the boundary or a prism with index $n_p > n_1$ may be clamped into the film in order to couple light in or out. A pair of input-output couplers clamped to a waveguide surface is illustrated in Fig.7.

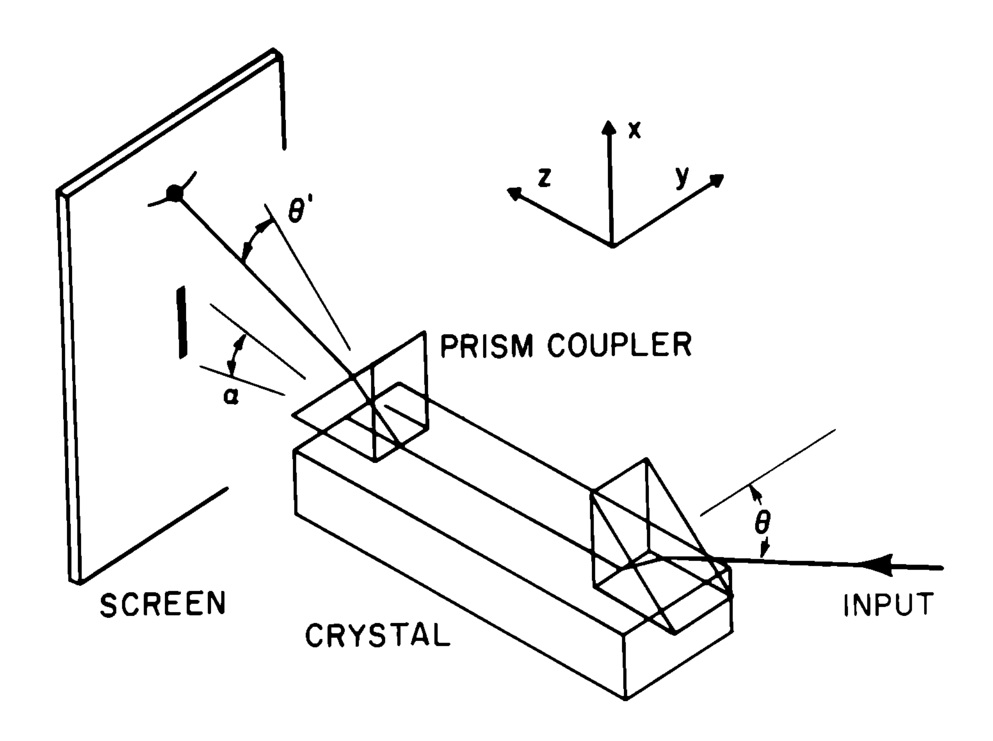


Figure 7 :
Input and Output Prism Couplers on a Waveguiding Crystal.

The wavefunctions for the guided modes can be obtained by inserting the refractive index profile into the wave equation. The results are analogous to those found in quantum mechanics, where index levels replace energy levels (see Fig. 8). The slab guide can be solved exactly but the graded guide usually requires approximate methods. However, the exponential profile.

$$\Delta n(x) = a_1 \exp(-x/b_1)$$

which may be used to approximate real profiles, may be solved exactly. [8] The number of modes M for slab and exponential guides are proportional to $(B/\lambda)(n_\infty A)^{1/2}$ and $(b/\lambda)(n_\infty a)^{1/2}$, respectively, for M large.

The propagation constants β for an exponential guide are plotted in Fig.9.[9] As with the asymmetrical slab guide there is a finite cutoff value for ξ, the normalized frequency. At cutoff, $\beta/k=n_\infty$ and far from cutoff, $\beta/k = n_\infty + a_1$.

Methods for providing waveguiding layers by plasma, Schottky barrier and p-n junction effects are summarized in Fig 10. [4]

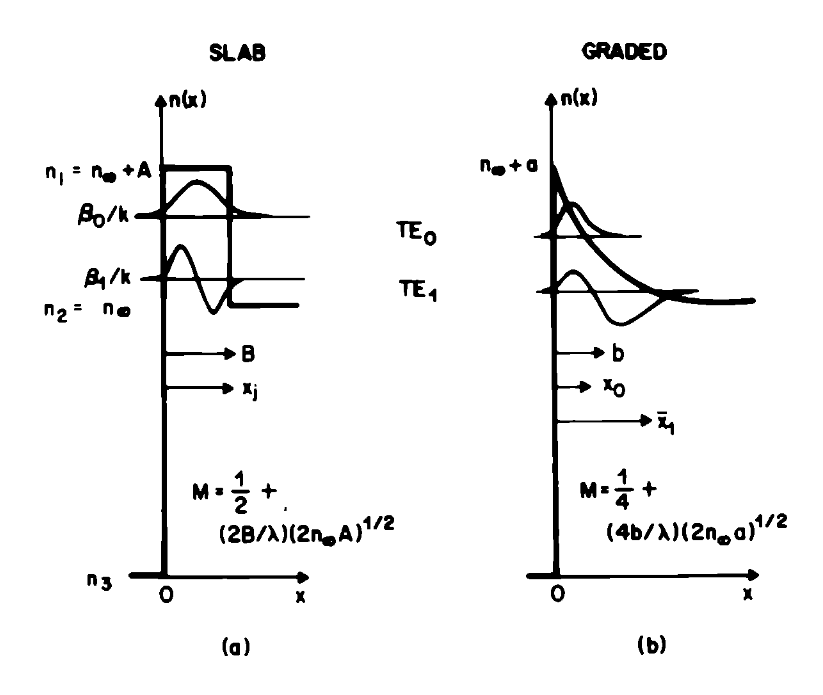


Figure 8 :
Index Profiles and Wavefunctions for Slab and Graded Guides.

Heteroepitaxial methods are summarized in Fig. 11.[4] And diffusion methods are summarized in Fig. 12. [4] In the 1-10μm range, GaAs and CdTe are among the best electrooptic materials. In the 0.4 to 4 μm range, $LiNbO_3$ is among the best electro- and acousto-optic materials. The magnetooptic garnets operate only near 1 μm (∿1.5-1.0 μm).

Most epitaxial semiconductor materials have too low resistivity to apply modulating voltages. However, a back-biased p-n junction overcomes the problem. An AlGaAs heterojunction diode modulator is illustrated in Fig. 13. At 1 μm, such a modulator produces 1 radian of modulation for 0.1 mW/MHz. [10]

An efficient thin-film magnetooptic garnet modulator has been demonstrated at 1.15 μm [11]

The refractive indices of several 2-6 compounds, which are electro - and acoustooptic, are plotted in Fig. 14. Optical waveguides have been formed by diffusing Se vapor into a ZnS substrate, for example, to produce a graded high index waveguide layer of Zn $Se_\alpha S_{1-\alpha}$. [12]

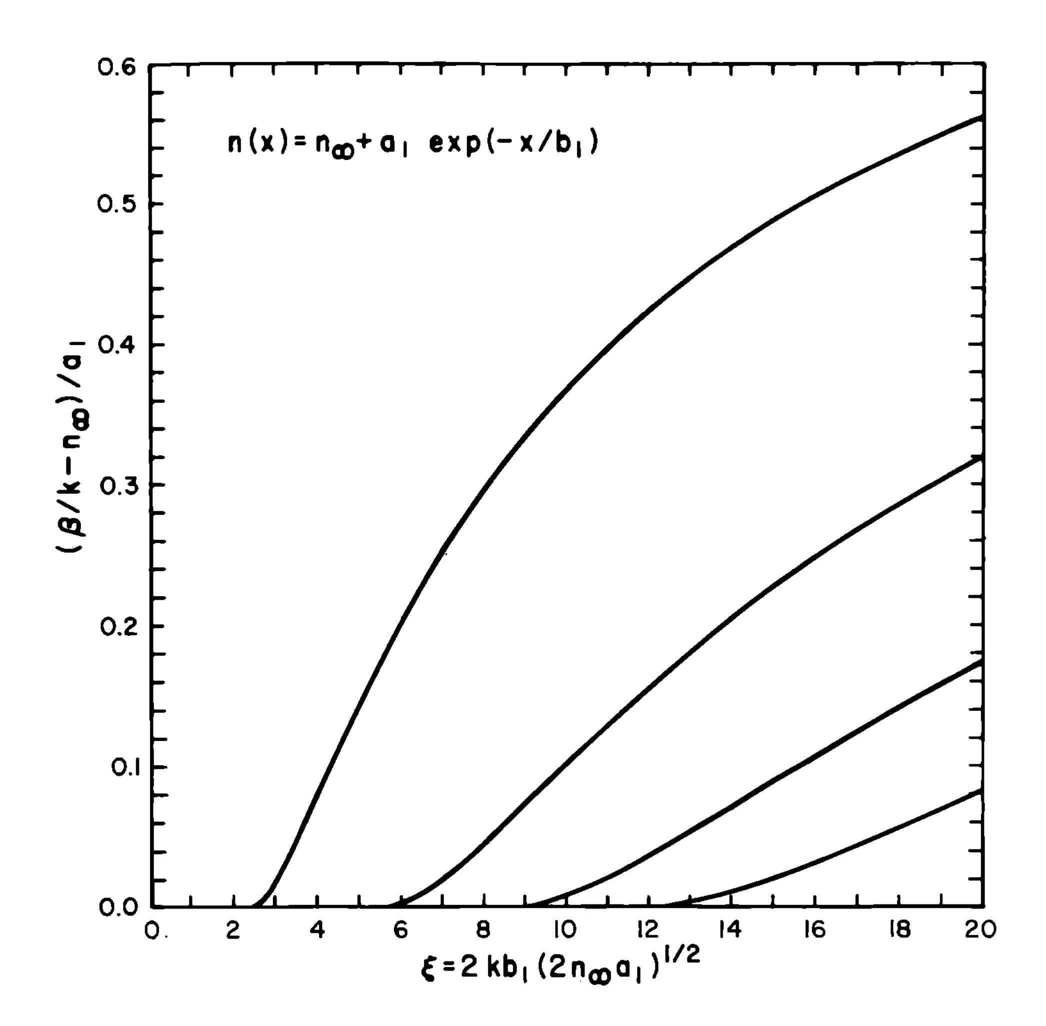


Figure 9 :
Propagation Constants for TE or TM Modes of Exponential Guide, for $n_\infty \gg n_3$.

PLASMA, SCHOTTKY BARRIER, p-n JUNCTION

$$\begin{cases} \Delta K = 2n\Delta n = -\dfrac{\omega_p^2}{\omega^2 + i\omega\tau^{-1}} \\ \omega_p^2 = \dfrac{Ne^2}{E_0 K m^*} \end{cases}$$

A. PROTON BOMBARDMENT

THIN-FILM MODULATOR MATERIALS

B. EPITAXIAL LAYER

C. DEPLETION AT JUNCTION

D. CARRIER INJECTION

Figure 10 :
Thin Film Modulator Materials : Plasma, Schottky Barrier , p-n junction. [4]

HETEROEPITAXIAL LAYERS

A. $(NH_4)_a K_{1-a} H_2PO_4$ ON KH_2PO_4

[$ADP_a KDP_{1-a}$ ON KDP]

B. ZnO OR AℓN ON Al_2O_3

C. GaAs-$A\ell_a Ga_{1-a}$As SYSTEM
EPILAYER AND DH p-n JUNCTION

D. GaP-$A\ell_a Ga_{1-a}$P

THIN-FILM MODULATOR MATERIALS

E. $LiNbO_3$ ON $LiTaO_3$

F. $Bi_{12}GaO_{20}$ OR $Bi_{12}TiO_{20}$ ON $Bi_{12}GeO_{20}$

G. MAGNETIC GARNET FILMS
[$R_3B_5O_{12}$]

Figure 11 :
Thin Film Modulator Materials : Heteroepitaxial Layers. [4]

DIFFUSED WAVEGUIDES

$$\begin{cases} \Delta n(x) = a_1 \exp(-x/b_1) \\ \Delta n(x) = a_2 \operatorname{erfc}(x/b_2) \\ \Delta n(x) = a_3 \sqrt{\pi}\, \operatorname{ierfc}(x/b_3) \\ \Delta n(x) = a_4 \exp(-x^2/b_4^2) \end{cases}$$

THIN-FILM MODULATOR MATERIALS

A. Cd(S,Se); Z(S,Se); (Zn,Cd)Se; (Zn,Cd)S

B. OUTDIFFUSED $LiNbO_3$ AND $LiTaO_3$

C. $LiNb_aTa_{1-a}O_3$

D. METAL - DIFFUSED $LiNbO_3$

Figure 12 :
Thin Film Modulator Materials: Diffused Waveguides.

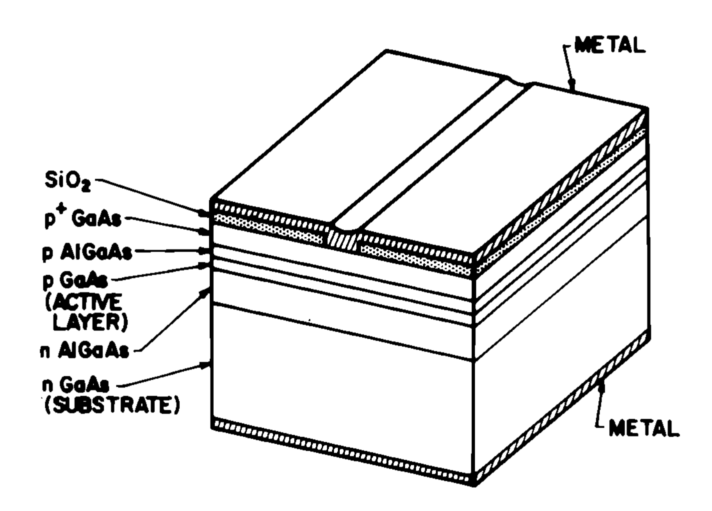


Figure 13 :
AlGaAs Heterojunction Diode Modulator. [10]

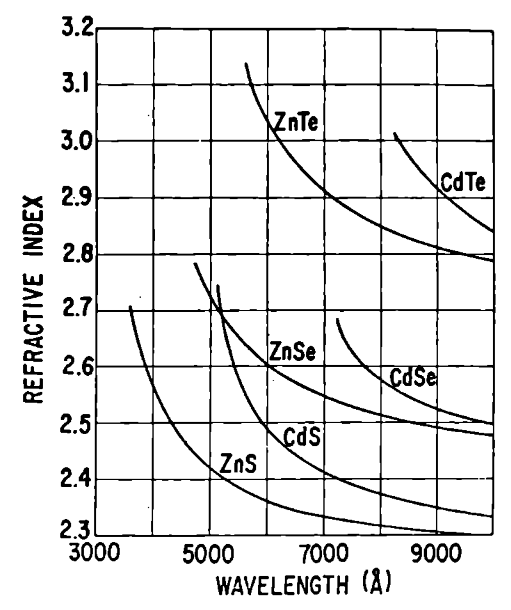


Figure 14 :
Refractive Indices of 2-6 Compounds. [12]

The refractive indices of $LiNbO_3$ are plotted against the stoichiometry index ν in Fig. 15. The extraordinary index increases as the Li/Nb ratio is reduced. Optical waveguides have been

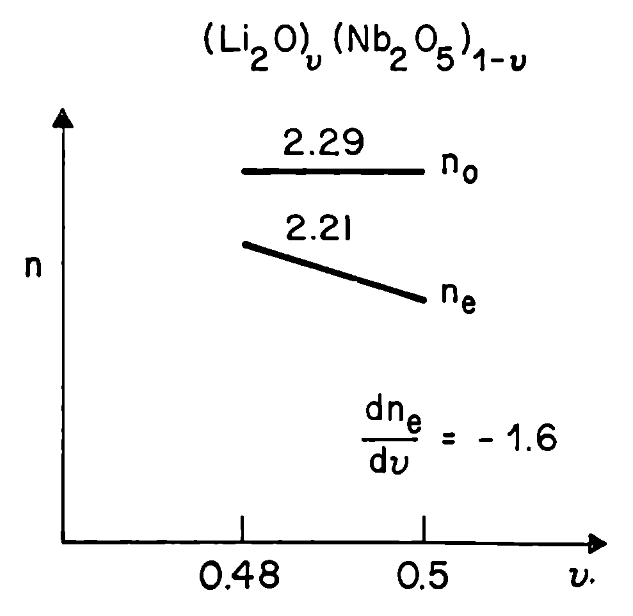


Figure 15 :
Refractive Index vs. Stoichiometry for $LiNbO_3$

made in $LiNbO_3$ and $LiTaO_3$ by heating the crystals in vacuum to diffuse Li from the surface. [13] Figure 16 shows the refractive index

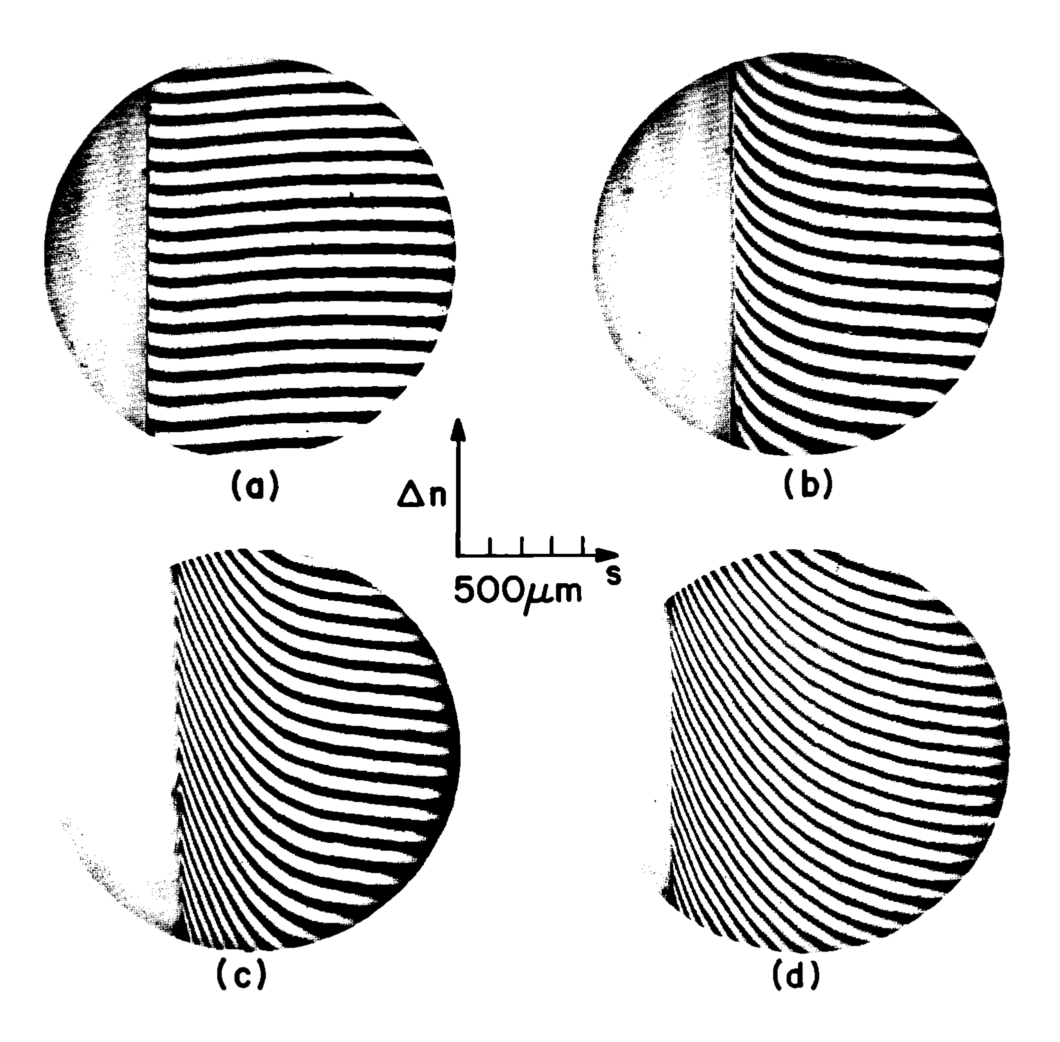


Figure 16 :
Interferograms of Outdiffused $LiNbO_3$ Crystals in (a) Ordinary and (b) - (d) Extraordinary Light. [13] .

profiles, obtained with the interference microscope, at the surface of outdiffused $LiNbO_3$ crystals for (a) the ordinary and (b) - (d) the extraordinary index. An electrooptic phase modulator formed by evaporating electrodes on the surface (Fig. 17) required 8 volts and 0.4mW/MHz to produce 1 radian. [13]

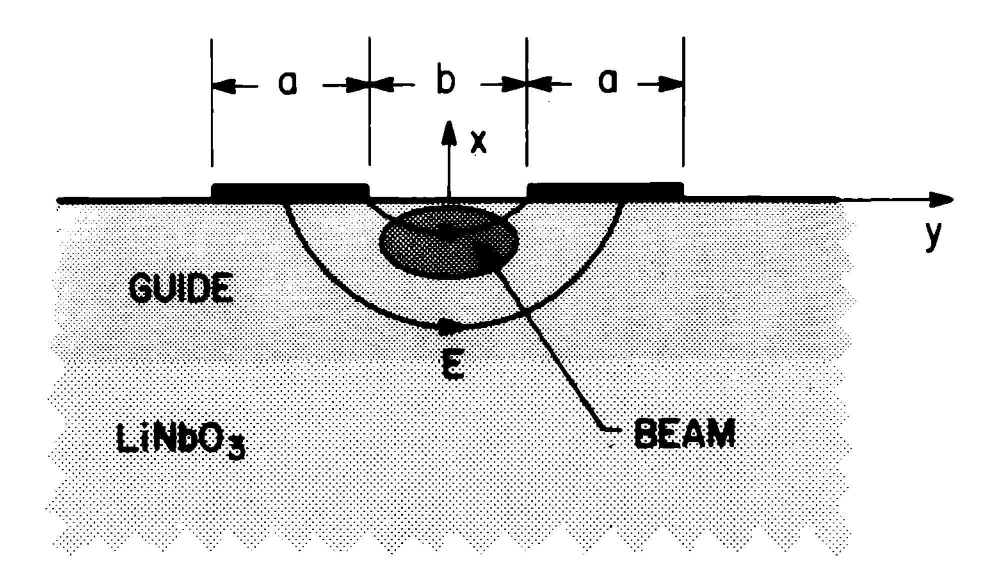


Figure 17 :
Coplanar Electrodes on Outdiffused $LiNbO_3$ Guide. [13]

$LiNbO_3$ has a higher refractive index than $LiTaO_3$. Waveguides have been formed on $LiTaO_3$ by melting and recrystallizing a film of $LiNbO_3$ on the surface. [14,15] Another method [16] is to evaporate metallic Nb on the surface of $LiTaO_3$ and heat the crystal in order to diffuse in Nb and form a graded guide of $LiNb_{\alpha}Ta_{1-\alpha}O_3$ It has also been found that many other metals may be similarly diffused into $LiNbO_3$ to form waveguides. [17]

In order to achieve amplitude modulation or beam deflection, a surface grating may be produced electro- or acoustooptically. A thin grating produces Raman-Nath diffraction while a thick grating produces Bragg diffraction with only one diffracted beam (see Fig 18). Since the acoustic field from a surface acoustic wave penetrates less than an acoustic wavelength Λ , and since the fringing electric field from a surface interdigital electrode penetrates less than the electrode period Λ, it is important that the optical field be confined closer then Λ to the surface. For the metal-diffused guides, the guide width is $\sim 1\ \mu m$.

One may also use an acoustooptic grating to couple a guided mode out of the waveguide into the substrate at a variable angle to produce a beam scanner. [18]

In order to realize the full advantage of a waveguide device, the beam must be confined in both transverse dimensions. This has been accomplished in an outdiffused $LiNbO_3$ planar guide by

etching away all but a narrow strip (19 μm wide × 8 μm high × 11mm long) as shown in Fig 19. Metal electrodes are evaporated on the sides of the ridge to provide the modulating field that interacts with the guided mode. With this device, 1.2 volts and 20 μW/MHz

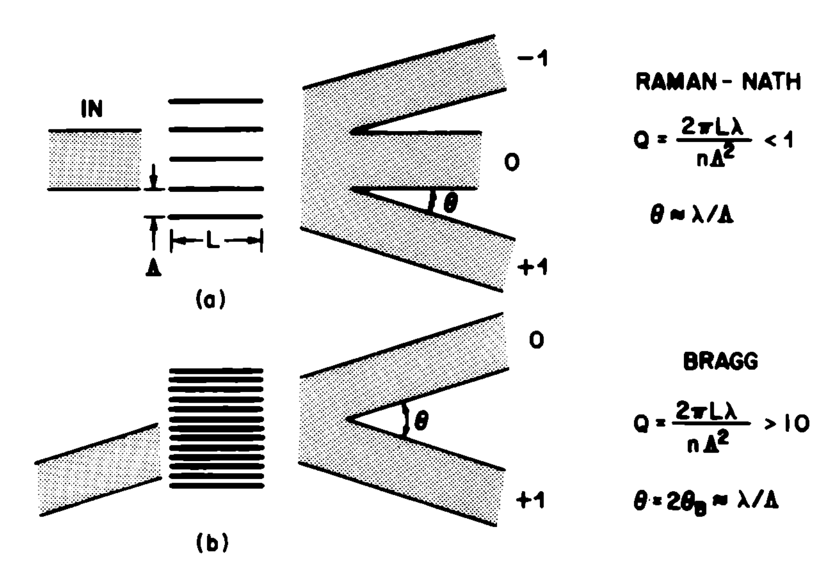


Figure 18 :
Raman-Nath and Bragg Diffraction.

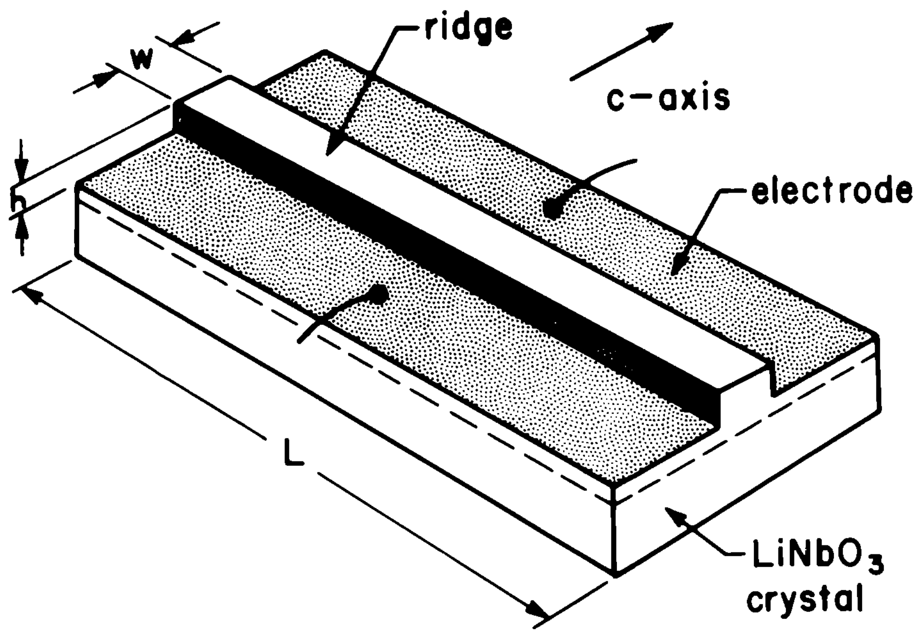


Figure 19 :
$LiNbO_3$ Ridge Waveguide Modulator. [19]

produce 1 radian of phase modulation. [19]

Phase modulation may be converted to amplitude modulation by phase grating diffraction as previously noted. Alternatively, heterodyne detection may be employed. Still another possibility is to insert the phase modulation in one or both arms of a balanced bridge (or interferometer) as shown in Fig. 20(A). Then variations

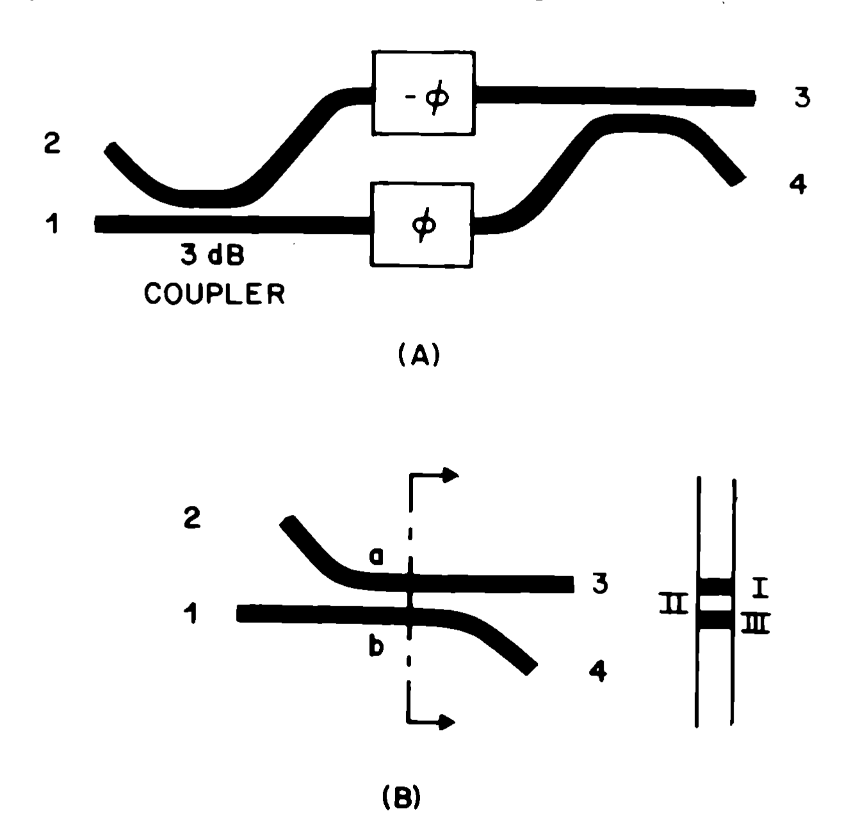


Figure 20 :
(A) Balanced Bridge Modulator/Switch.
(B) Directional Coupler Modulator. [20]

of ϕ by $\pi/2$ switch the output between ports 3 and 4. The 3dB couplers may be external beam-splitters or, one day, integrated optical waveguide directional couplers. Another proposal [20] combines both the phase shift and coupling functions as in Fig. 20 (B).

IV. SUMMARY

We have seen that waveguide optical devices offer better power efficiency ($\sim$20 μW/MHz) than bulk devices ; and, further , that these devices may be less difficult and expensive to fabricate.

■ REFERENCES

[1] I.P. Kaminow and E.H. Turner, "Electrooptic Light Modulators", Proc. I.E.E.E. , 54 , 1374 (1966)

[2] I.P. Kaminow, An Introduction to Electrooptic Devices, Academic Press, N.Y., (1974).

[3] F.S. Chen, "Modulators for Optical Communications" Proc. I.E.E.E., 58 , 1440 (1970).

[4] I.P. Kaminow, "Optical Waveguide Modulators", I.E.E.E. Trans. on Micr. Theor. and Tech., January 1975.

[5] G. White and G.M. Chin, "Traveling Wave Electrooptic Modulators", Optics Commun. 5, 374 (1972)

[6] F.S. Chen and W.W. Benson, "A Lithium Niobate Light Modulator for Optical Communications", Proc. I.E.E.E. , 62 , 133, (1974)

[7] A.W. Warner and D.A. Pinnow, "Miniature Acoustooptic Modulators for Optical Communications", I.E.E.E. J. Quant. Electr. 9 , 1155 (1973).

[8] E.M. Conwell, "Modes in Optical Waveguides Formed by Diffusion", Appl. Phys. Letters 23 , 328 (1973).

[9] J.R. Carruthers, I.P. Kaminow, and L.W. Stulz, "Outdiffusion-Kinetics and Optical Waveguiding Properties of Outdiffused Layers in Lithium Niobate and Lithium Tantalate", Appl. Optics 13 , (1974).

[10] F.K. Reinhart and B.I. Miller, "Efficient GaAs-$Al_xGa_{1-x}As$ Double-Heterostructure Light Modulators", Appl. Phys. Letters 20 , 36 (1972).

[11] P.K. Tien, R.J. Martin, R. Wolfe, R.C. LeCraw and S.L. Blank, "Switching and Modulation of Light in Magnetooptic Waveguides of Garnet Films", Appl. Phys. Letters 21 394 (1972).

[12] W.E. Martin and D.B. Hall, "Optical Waveguides by Diffusion in II-VI Compounds" , Appl. Phys. Letters 21 325 (1972).

[13] I.P. Kaminow, J.R. Carruthers, E.H. Turner and L.W. Stulz, "Thin-Film $LiNbO_3$ Electrooptic Light Modulator", Appl. Phys. Letters , 22 540 (1973).

[14] S. Miyazawa, "Growth of $LiNbO_3$ Single-Crystal Film for Optical Waveguides", Appl. Phys. Letters 23 198 (1973).

[14] S. Fukunishi, N. Uchida, S. Miyazawa and J. Noda "Electrooptic Modulation of Optical Guided Wave in $LiNbO_3$ Thin Film Fabricated by EGM Method", Appl. Phys. Letters, 24 , 424 (1974).

[15] P.K. Tien, S. Riva-Sanseverino, R.J. Martin, A.A. Ballman and H. Brown, "Optical Waveguide Modes in Single Crystalline $LiNbO_3$-$LiTaO_3$ Solid Solution Films", Appl. Phys. Letters 24, 503 (1974).

[16] J.M. Hammer and W. Phillips, "Low-Loss Single-Mode Optical Waveguides and Efficient High-Speed Modulators at $LiNb_xTa_{1-x}O_3$ on $LiTaO_3$ ", Appl. Phys. Leters 24, 595 (1974).

[17] R.V. Schmidt and I.P. Kaminow , "Metal-Diffused Optical Waveguides in $LiNbO_3$ ", Appl. Phys. Letters 24 ,(1974).

[18] F.R. Gfeller and C.W. Pitt, "Colinear Acoustooptic Deflection in Thin Films" Electron. Letters 8 549 (1972) F.R.Gfeller , "Acoustooptic Scanner ", Integrated Optics Meeting, New Orleans (1974).

[19] I.P. Kaminow, V. Ramaswamy, R.V. Schmidt and E.H. Turner, "Lithium Niobate Ridge Waveguide Modulator", Appl. Phys. Letters, 24 (1973).

[20] S. Kurazono, K. Iwasaki, N. Kumagi, "A New Optical Modulator Consiting of Coupled Optical Waveguides", Electr. and Comm. in Japan 55C , 103 (1972).

INTEGRATED OPTICS : A PREJUDICED PERSPECTIVE

D.B. Ostrowsky

Thomson-CSF Domaine de Corbeville

BP n°10 91401 Orsay France

I. INTRODUCTION

The advent of the laser, stimulated, in the early 1960's a great interest in optical communication systems. After a period of rather intense interest spanning several years the field quieted down as research workers began to appreciate the enormous difficulties involved. The existing optical and electro-optical elements did not permit the assemblage of reliable, performant, terminal systems, and, there was no suitable transmission medium since free atmospheric transmission was found to be too unreliable and the existing optical fibers had losses on the order of 1000 db/Km.

Towards the end of the decade, however, a number of developments combined to restimulate a very broad based interest in this field. These were ;

a) The announcement [1] of fibers having losses on the order of 20 db/km
b) The realisation of room temperature CW GaAs lasers [2]
c) The appearance of a number of theoretical and experimental papers which indicated the possibility of realizing a family of optical and electro-optical devices in a thin film form, which gave rise to the new field known as " Integrated Optics" which will be the object of this paper [3,4].

In great part due to the overall interest in optical communications created by a) and b) integrated optics attracted a impressive number of research workers in many countries. At the Topical Meeting on Integrated Optics held at Las Vegas in Feb. 1972, approximately 55 papers were presented. Two years later, when the Topical Meeting was held at New Orleans about 75 papers were delivered, but, in my opinion, the majority of participants left

convinced that the advances made since 1972 were technical rather than inventive. This suggests that we may have reached a sort of plateau, philosophically similar to the one reached in the mid-sixties by engineers working on optical communication systems : the early glow of unbridled enthusiasm has worn off as the difficulties become more evident. In any case a reunion such as Cadarache presents a good occasion for attempting to realise a sort of stock taking discusion since there are people present who are rather expert in each of the related domains.

In an attempt to begin such a discussion I will present, after a brief general introduction to integrated optics, the point of view which has evolved, among the workers in our own laboratory, concerning the current state of the art of integrated optics. Rather than attempt to present a comprehensive review article I shall begin by discussing the two fundamental technological building blocks underlying integrated optics : materials and microstructures. This will lead into a somewhat more detailed discussion of two integrated optical components that have not yet been realized but which, we feel, present reasonable goals for the medium range future (1 or 2 years). These are : the electrically switched directional coupler, which we call a COBRA (from the French : Commutateur Optique Binaire Rapide) and a frequency doubled GaAs laser.

▪ II. INTEGRATED OPTICS : A BRIEF INTRODUCTION

In order to take full advantage of the possibilities offered by coherent optical communication systems it will be necessary to develop a family of optical and electrooptical components permitting the processing of optical signals. It has been suggested [3, 5] that such components could be fabricated in a planar form using techniques derived from those developed for the production of integrated electrical circuits and the name "integrated optics" has been adopted for this type of effort.

While the underlying idea had already been suggested by D.B. Anderson in 1965 [5] the real foundation was laid with the appearance of the Sept. 1969 edition of the Bell System Technical Journal. In an introductory article Miller [3] outlined the possibilities and suggested the form various components might take. To get an idea of what was accomplished to date lets look at Fig. 1 which reproduces the suggestions of Miller and states whether the element has been realized.

We see that five out of the six have been realized. Nevertheless, we must recognize that most of these realisations have been somewhat acrobatic prototypes , rather far removed from what will be ultimately needed. This should become clear when I discuss the current state of the art concerning the fundamental building blocks of these components : materials and microstructure fabrication.

In fact, due to the difficulty of the microstructure problem, the vast majority of the experiments performed in integrated optics to date, have used a planar waveguide geometry (Fig 2a), rather than laterally bounded structures (Fig 2b,c). I shall therefore,

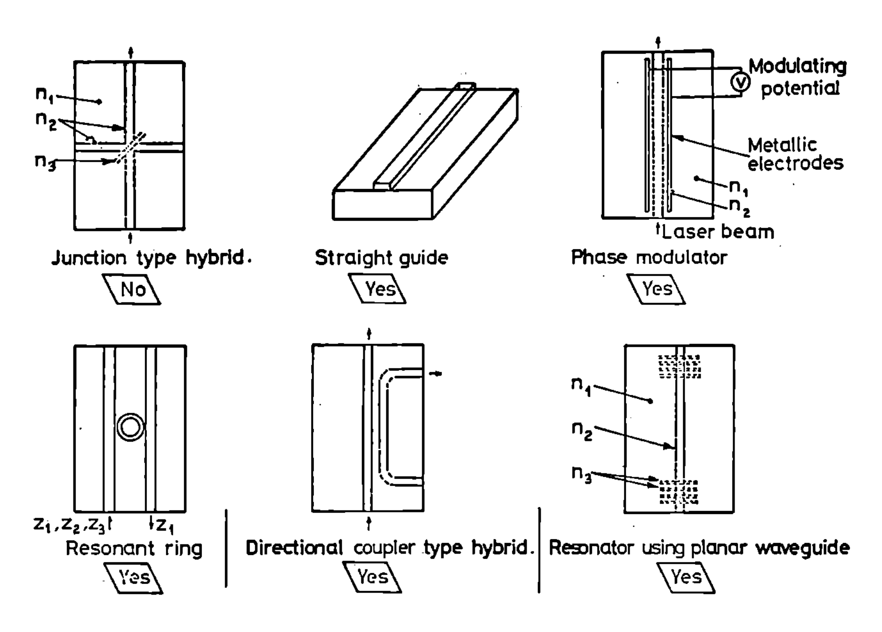
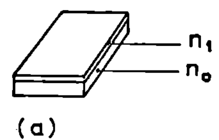


Figure 1 :
Some suggested elements and whether they've been realized.

Planar guide



(a)

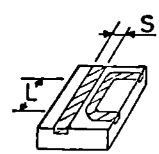
Straight guide

n2
n1
w
h

n_1 = 1.5, n_2 = 1.515
w = 5 microns
h = 1 micron
Δw (loss < 1dB/cm) = 05micron

(b)

Directional coupler



Coupler : S = 1 micron
L = 3.9mm
ΔS (L ± 20% = 0.1 micron

(c)

Figure 2 :
Some configurations and their tolerances. Typical parameters for the planar guide are : n substrate=I.5, n guide=I.6I thickness I μ

briefly describe the phenomena of waveguiding in such films.

While several authors [6,7] have presented analyses applicable to an arbitrary number of superposed films I shall quickly outline here a simple way of treating a single layer waveguide.

A) Ray propagation

A ray diagram of light propagating in such a waveguide is given in Fig 3.

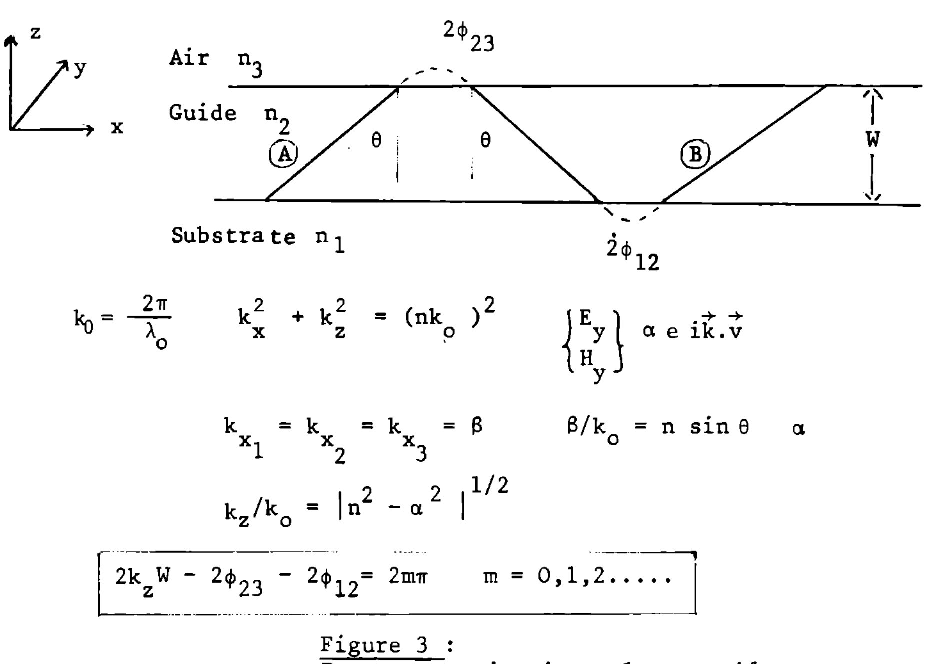


Figure 3 :
Ray propagation in a planar guide : definitions and the resonance condition

The ray can be described in terms of its k vector which points in the direction of ray propagation with :

$$|k_i| = n_i |k_o|$$

$$k_o = \frac{2\pi}{\lambda_0} \qquad \lambda_0 = \text{vacuum wavelength}$$

$$k_{x_i} = k_i \sin\theta_i \qquad k_{z_i} = k_i \cos\theta_i$$

While the angle of reflection (θ_1) equals the angles of incidence (θ_1) one must take into account phase delays introduced when the ray undergoes total reflection at the boundaries. These delays (ϕ's) account for the fact that the ray "spends some time"

outside of the guide. This is the way one "considers" evanescent waves when using a ray picture of light propagation. We will discuss these waves in more detail in the Field Distribution section.

B. Resonance

Using this ray picture we can now determine the resonance condition which allows propagation within such a guide. For a resonance to occur we must have constructive interference between the waves propagating in the + and -z directions. This is equivalent to demanding that in one round trip ((A) (B)) the phase of the wave is increased by an integral number of 2π radians. That is we demand :

$$2k_{z_1} W + 2\phi_{23} + 2\phi_{12} = 2m\pi \ , = 0, 1, 2 \ldots\ldots$$

Since they are exponential functions of the ϕ, θ, and the various indices of refraction this is a transcendantal equation which must be solved graphically or with a computer. Doing this one finds, for a given set of indices, a curve of guide thickness, W versus n sin θ (which is the effective index of refraction of the guiding structure of the form given in Figure 4.

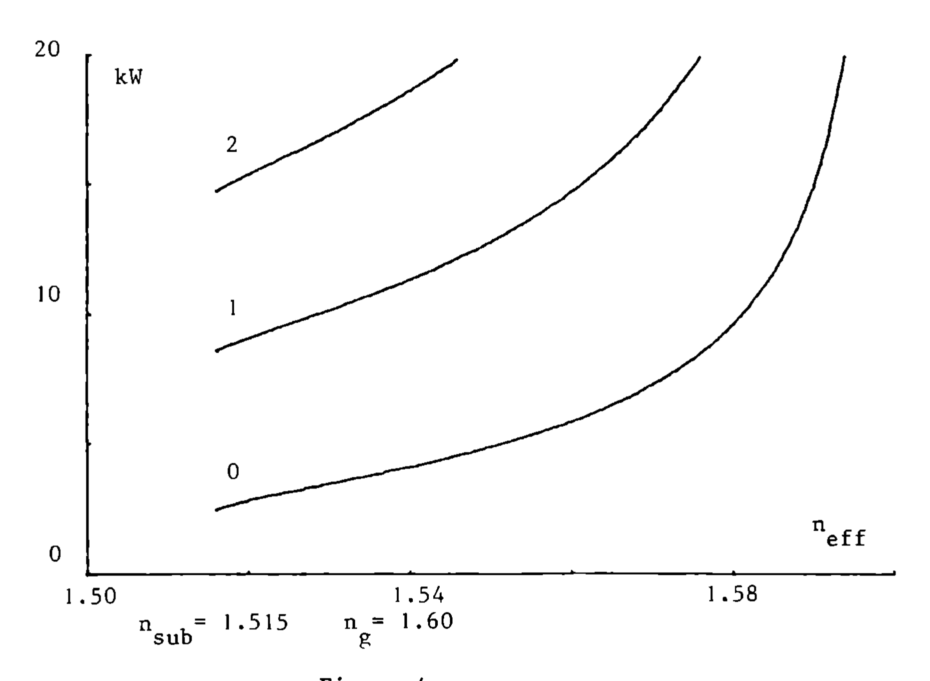


Figure 4 :
Effective index versus guide thickness

C. Field Distribution

For a given mode, once one has determined the θ associated with the mode, one has also determined the form of the field distribution, describing the optical electric field, since the field is always of the form :

$$E \propto e^{i(k_x x + k_z z)}$$

and $k_x = k_0 \sin\theta$, $k_z = k \cos\theta$

Although the mathematical form is always the same, the configuration of the field will be quite different in the three different regions since k_z will be *imaginary* in regions 0 and 2 due to the condition of total reflection.

Since, from Maxwell equations we have, for each region :

$$(kx_i)^2 + (k_{z_i})^2 = (n_i k_0)^2$$

when k_{x_i} is bigger than $n_i k_0$, as is the case in regions 0 and 2, k_z *must be imaginary*. The wave in these regions is of the form :

$$E \propto e^{i(k_x x)} e^{-k_z z}$$

The wave amplitude decreases exponentially as one moves away from the guide boundary. This type of wave is called an *evanescent wave*. A graph of the field distribution in all three regions, for the modes of the guide described by Fig. (2) is given in Fig. 5.

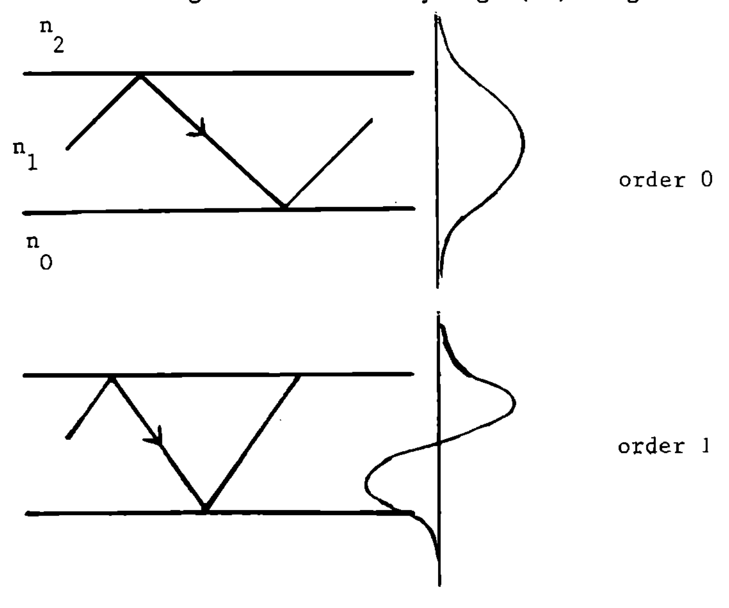


Figure 5 :
Field amplitude distribution

The main results of this analysis are the following :
Energy can propagate in the guides in certain discrete modes and with each permitted mode we can associate an angle of propagation and an effective index of refraction.
A portion of the total guided energy is carried by evanescent waves which lie outside the guiding film.

This is important since it permits interaction with materials lying outside the guiding film.

D. Coupling

Since Dr. Kogelnik has discussed this extensively I shall simply recall the usual techniques :
The two techniques which have proven to be the most useful, to date , are shown in Fig 6 : the *prism coupler* and the *grating coupler*. The prism coupler, which was used by Tien, Ulrich and Martin in their first demonstration (1969) of light guiding in thin films works as follows.

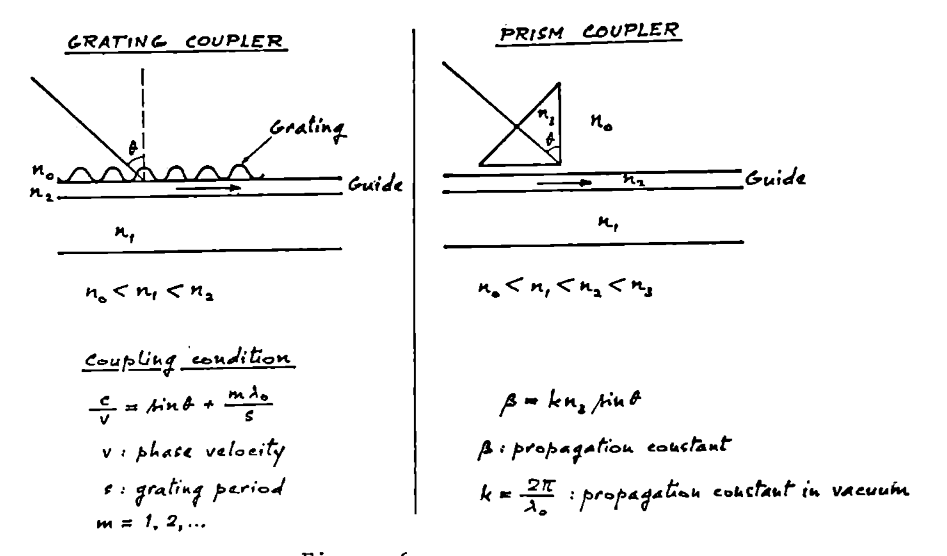


Figure 6 :
a) Grating coupler
b) Prism coupler.

A high index of refraction prism is pushed tightly against the upper surface of the guide. A beam of light enters the prism and undergoes total reflection at the base since there is always a low index of refraction, layer of air between the prism and the guide. This total reflection creates, *an evanescent wave* which can traverse the air layer and enter the guide, if this wave has the same "effective index of refraction" as a mode of the guide, that is, if $n_p \sin\theta_p = n_g \sin\theta_g$, there will be a resonant transfer of energy into the guide. Therefore by varying the angle of

incidence in the prism one can attain the resonance and excite the guide. While with an appropriately tapered prism 100% coupling efficiencies are possible a typical value, relatively easy to actually obtain with a simple right angle prism, is on the order of 30%.

Coupling out of the guide can be accomplished by the inverse process since the evanescent portion of the guided wave couples into the prism and excites the outgoing wave. In this direction we can always obtain 100% coupling since all the light will, eventually be coupled out.

While the prism coupler is an extremely handy experimental device it is not really compatible with our basic concept of integrated optics : that of fabricating all the necessary elements in a thin film geometry. In 1972, however, research workers at IBM proposed a planar coupler : the grating coupler Fig. 6a, by depositing a photosensitive resine upon the guide and exposing the resine to the interference fringes created by interfering laser beams (i.e. holography).

The grating coupler is the optical analog of a periodic antenna, which is used in the microwave domain. A beam is incident upon the surface so that one of the diffracted orders is in resonance with the guided wave (see coupling condition, Fig 6a).

Both the prism and grating coupler allow us to couple freely propagating modes to integrated optical elements. They do not resolve a more fundamental problem : that of coupling I O elements to mono-mode fibers. This problem certainly deserves more attention than it has received to date.

Before leaving the subject of coupling we should note that even modest coupling efficiencies can provide rather high light power densities in the guides. If one has a 5 mW laser with a 1 mm diameter beam one has a power density of 0.5 Watts/cm^2. If we couple 50% of this power into a guide having a 10 microns x 5000 Å section we obtain a power density of about 50000 W/cm^2. This power concentration possibility will be important for high power applications such as non-linear optics (frequency conversion, etc...).

Having completed our introductory remarks , we'll now consider the choice of materials for integrated optics.

▪ III. MATERIALS

In the years following the original experiment on light guiding in thin films [4], carried out using a polycrystalline ZnO film deposited on glass, research workers throughout the world carried out an extensive search for materials susceptible to being useful for integrated optics. An enormous number of experiments were performed trying practically all imaginable types of films (organic, inorganic, amorphous, crystalline, polycrystalline, etc...) using practically all known methods of film fabrication (evaporation , sputtering, spin-coating, epitaxy, implantation, diffusion , etc.). After fabrication one attempted to couple light into

and out of the films to determine the propagation losses (due to absorption and scattering). Since integrated optical structures will typically be several mms long, and we'll wish to connect a number of these elements it appears that guides will be useful when their losses are below about 0,5 db/cm. This corresponds to a loss of about 10% of the guided energy in 1cm of propagation.

Examining Table I we see that there are a fair number of materials that can be used to form passive waveguiding structures.

	TYPES	Fabrication Techniques	Index	Typical loss
Amorph	Glass	Sputtering, Ion exchange Implantation, diffusion	1.46-1.75	1db/cm(6328Å)
	Ta_2O_5	Sputtering, oxidation	2.2	1db/cm(6328Å)
Monocristalline	ZnO/Al_2O_3	Heteroepitaxy	2.2	2db/cm(6328Å)
	$(Li_2O)(Nb_2O_5)_{1-x}$	Exodiffusion	2.214-2.22	1db/cm(6328Å)
	$LiNb_xTa_{1-x}O_3$	Diffusion	2.18-2.237	1db/cm(6328Å)
	GaAs	Heteroepitaxy implantation	3.6	5-10db/cm(1.15μ)
	Garnet	Heteroepitaxy	1.94	1 3db/cm(1.15μ)
Organic	VTMS	rf discharge polymerization	1.531	0.04db/cm(6328Å)
	Shipley AZ 1350	Spin coating	1.6	2db/cm (1.06μ)

TABLE I

The necessary technological control of various film parameters have yet to be delineated and demonstrated. Nevertheless specialists in thin fabrication technology indicate that sufficient control ($\Delta E \approx 100$ Å , $\Delta n \approx 10^{-3}$) is within reach using existing techniques. The choice of material will probably be dictated by the tolerable index range and ease of fabrication for specific problems.

In the domain of active materials the situation is more complicated, since these materials will have to perform many diverse functions. Among the more obvious are : amplification, modulation, isolation and detection. The first question we must examine is : do we need one universal material for the realization of these functions.

My own response is no for the forseeable future. If we examine the best candidate for such a universal material - GaAs and its family - one finds some important drawbacks. GaAs guides are not

particularly good, at 1,15μ , 5 to 10 db/cm is common, and is even worse , of course , at 0,8 - 0,9 μ . Its interest as a material for electrooptic modulation is largely due to the fact that the double heterostructure permits the realization of very thin modulators [8]. Using equal geometries the factor of merit of GaAs as an electrooptic modulator is about 10 times lower than that of the $LiNbO_3$ / $LiTaO_3$ family. Excellent photodetectors can be realized in GaAs, [24] but, it is of course, non-magnetic.

For these reasons I feel that while GaAs continues to merit enormous interest it will not resolve all our problems. This suggests the following approach : let's use GaAs for what it does best - emission and probably detection of light (along with integrated electronics) - and see if we can't couple it to other types of materials for realizing integrated systems.

What are some of these other materials ?

For modulators and commutators among the most attractive are the members of the $LiNbO_3/LiTaO_3$ family which will certainly be heavily investigated in the coming years. Exodiffusion of $LiNbO_3$ and $LiTaO_3$ and Nb diffusion into $LiTaO_3$ have already permitted the realization of integrated optical modulators whose performance exceeds bulk devices by at least an order of magnitude [9,10]. This will be discussed in more detail in the lecture of Dr. Kaminow this afternoon. I will limit myself to discussing a bit later a specific component we're trying to realize in our laboratory, based on such materials - the COBRA (Commutateur Optique Binaire Rapide).

For the magnetic materials we'll need for non-reciprocal devices there is a great amount of material research to be done. While the Gadolinium and Scandium doped iron garnets are attractive materials in the near IR [11,12] there is a continuing interest in developing materials that could function at shorter wavelengths and allowing a rotation greater than the 200°/cm offered by these materials. Bismuth substitued garnets [13] offer a possibility of slightly extending the usable wavelength and increasing the rotation/cm by an order of magnitude and will certainly be investigated in the coming years. In any case, the possibilities of material combinations in this family is enormous and I refer you to the article of Mr. Désormière in this volume for further discussion of this point. I'd like to point out however in closing on this topic that an important piece of theoretical work remains to be done in this field.

To date, all the work in this field both, experimental and theoretical [14,15], has been concerned with planar guides. It will be important to extend this work to the more complicated case of laterally bounded guides.

In conclusion, I believe we still have a great amount of basic material research to do before we can begin to realistically design systems. Much of this work will certainly continue using planar guiding structures, but, I believe we will begin

to evaluate materials more critically in terms of their compatibility with microstructure fabrication techniques which are the subject of the following section.

IV. MICROSTRUCTURE - FABRICATION AND EXAMINATION

In this section we'll discuss the second fundamental problem facing integrated optics : microstructure fabrication and examination.

Some examples of the types of problems encountered are shown in Fig. 2b,c. Consider the problem of fabricating a simple laterally bounded guide such as that shown in Fig.2b. The structure will typically be several microns wide, less than 1 micron thick, and have a length ranging from millimeters to centimeters. If the index difference between the guide and the surrounding media is on the order of 1% the edge roughness must be held below 500 Å to have losses below 1db/cm [16].

In addition to this edge roughness requirement there is a dimensional tolerance problem caused by the fact that certain integrated optical circuit elements require 1000 Å (or better) precision over *macroscopic* (mm to cm) distances. Consider the directional coupler schematically shown in Fig 1.Since the coupling length (the length at which the energy will pass completely from one guide to the other) is proportional to e^{-c/k_c} where c is the distance between the guides and k_c is the component of the wave vector directed between the guides, small changes in c can strongly modify the coupling length. For a coupler having the geometry shown in Fig 2c. the coupling length, L, calculated using the approximation of Marcatili [17] will be 3.9 mm and $\frac{\Delta L}{\Delta c} \approx 7000$. A 1000 Å difference in the separation distance will change the coupling length by 20%. The fabrication of such a coupler having a prescribed coupling length requires, therefore, maintaining better than 1000 Å precision over several mm.

In 1969 Miller [3] suggested the use of electron-beam masking to overcome these difficulties and this suggestion has been followed at several laboratories, including our own. In collaboration with the techniques shown in Fig.7 this led to the production of fairly good guides (edge roughness below 500 Å , losses on the order of about 10 db/cm at 6328 Å) by the time of the Las Vegas Meeting [18].

In our own work we used a scanning electron microscope whose beam was controlled by the output of a flying spot scanner which reads a photographically produced mask. Although this provided reasonable flexibility we still could not realize designs having the width to length ratio of the directional coupler we have discussed due to the fact that electron beams can be controlled with the necessary precision over one millimeter at best. Furthermore one should not forget that some day we will have to integrate a large number of elements. We have therefore utilized an electron beam machine which interferometrically controls

substrate position to span the necessary macroscopic distances. The machine (developped by J. Trotel and his co-workers at our laboratory) and the program developped for optical waveguide fabrication is described in the following sections.

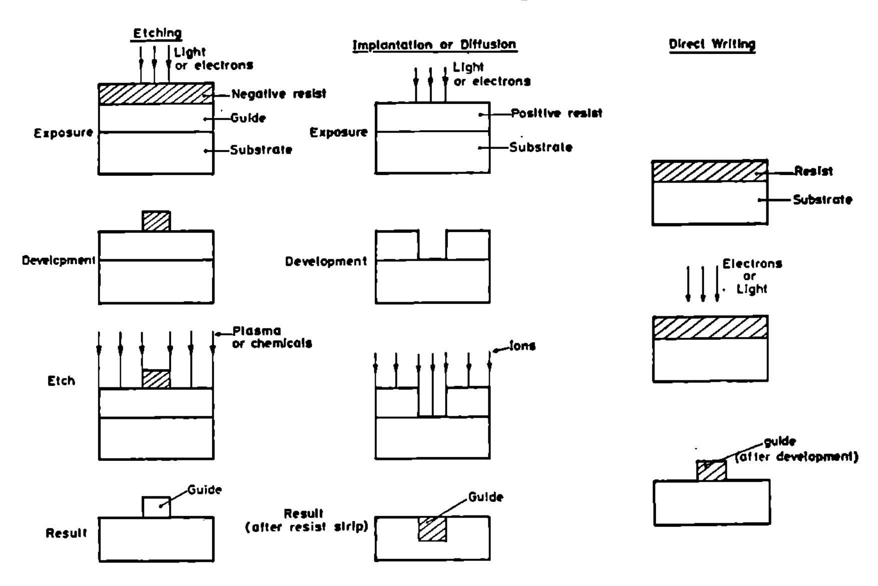


Figure 7 :
Schematized fabrication techniques

A) The masqueur electronique (ME)

The ME was designed to permit electron beam exposure over a 5x5cm surface with a 1000 Å precision between any two points.

A schematic of this machine is shown in Figure 8.

The essential elements are :

1) A scanning electron microscope system (CAMECA MEB/07),
2) A table allowing X-Y displacement of the sample being masked with an interferometric position control.
3) An automatic control system including :
   - A Varian 620/i and associated peripherals,
   - An elementary shape generator , all figures are composed of elementary rectangles, parallelograms or triangles formed by exposing a number of points,
   - Point seperation is usually 400 Å

To trace a given design the automatic controls procede as follows :

1) The X-Y motors drive the table to the designated origin with a precision of 2.5μ.
2) The interferometer measures the real position of the table with a precision of 400 Å .
3) The difference between the real and desired position is computed and the computer uses the compensating coils to displace

the electron beam to the desired position.
4) The desired figure is built up from the elementary ractangles in a 100 × 100 μ field, individual points being scanned at 500 KHz. During the writing process it is possible to visualize the figure being written since the masqueur is functioning as a scanning electron microscope.
5) The X-Y motors drive the table to the next desired position and the process is repeated.

In this way the desired pattern can be built up over a 5 × 5cm surface.

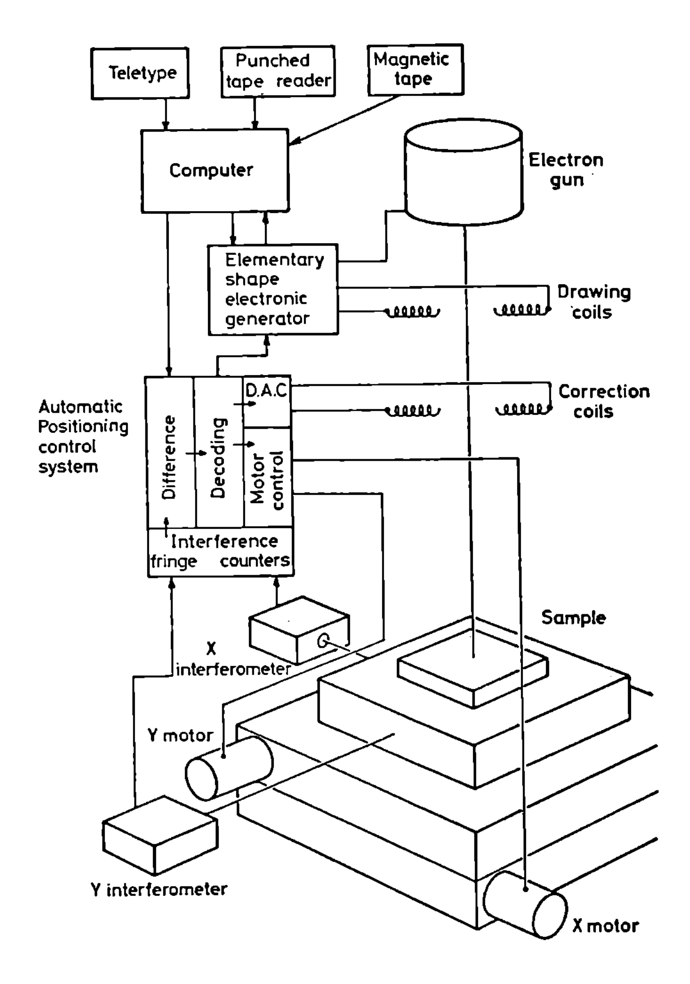


Figure 8 :
Masqueur schematic

B) The optical waveguide program : Trakus

Since the Masqueur was originally designed to trace step and repeat programs for electronic circuits it was necessary to deve-

lop a special program for waveguides, which we shall briefly describe.

The program TRAKUS enables us to form waveguides mask containing the following elements :

1) *Lines* : Horizontal or vertical lines are formed by repeating rectangles. The usual widths are 1,5 or 10μ with lengths ranging from mms to cms.
2) *Arcs of circles* : Arcs are formed by repeating a series of rings. For a 5 μ wide guide one needs to have ring centers separated by approximately 1μ in order to keep edge irregularities below 500 Å . Typical widths are again 1, 5 or 10μ with radii of curvature from 0,1 to 5 mm.
3) *Macroscopic squares* : These are built up of elementary rectangles and are used to facilitate coupling with an incident laser beam.
4) *Triangles* : These are built up from elementary squares and are used as transitions between macroscopic and microscopic guide sections. A typical triangle tapers from 25 or 50 μ to 1,5 or 10 μ.

To draw a desired figure one simply enters the positions and dimensions of the desired elements and the sequence to be followed while tracing composite design. TRAKUS then directs the tracing of the pattern.

In the following section we will discuss the fabrication and examination of actual waveguides fabricated using these techniques.

C. Experimental results : Fabrication and Examination

Using, in particular, the direct writing technique in PPS,[19] we have realized a number of family complex components. These include :

* straight and curved guides,
* directional couplers,
* triangular (or horn) transitions,
* resonant rings placed between straight guides

An example of light diffused out of a directional coupler is shown in Fig 9.

We believe that our results show that the pattern delineation problem is on its way to being solved. Even if e-beam writing is relatively slow, rapid duplication techniques [20,21] (e-beam, X ray lithography, and optical copying) seem capable of providing the necessary resolution.

While we have made extensive use of PPS due to the extreme simplicity of direct writing we have also realized complex structures in inorganic materials. Among them are glass (by ion etching using masks), fused quartz (proton implantation using gold masks) and $LiTaO_3$ (Niobium diffusion). The implantation and diffusion techniques have already yielded 5 micron wide guides with losses on the order of 1 db/cm. The etched guides have had 2-5 db/cm losses but this can be reduced by fabricating ridged or multi-layer structures.

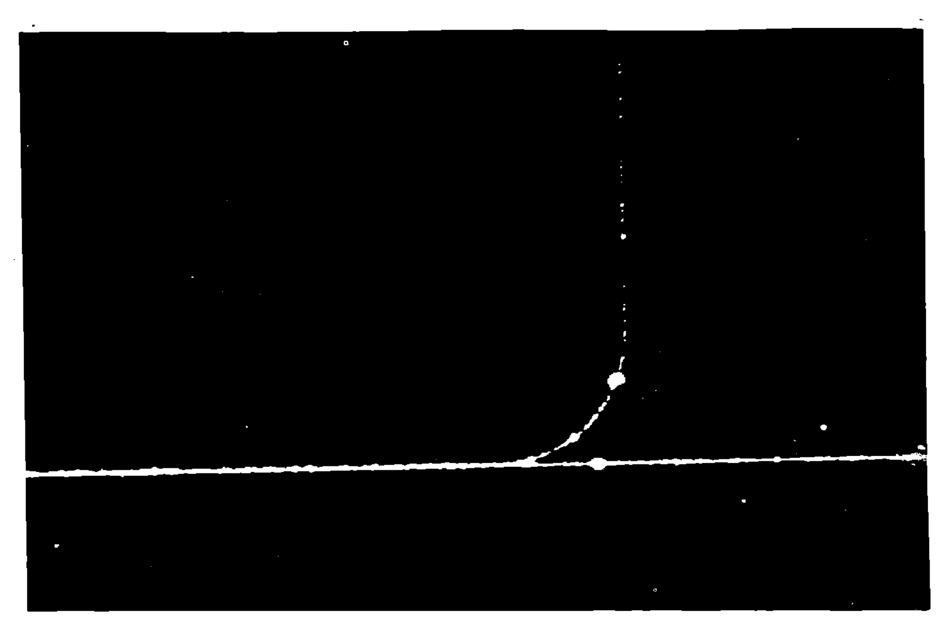


Figure 9 :
Output of a directional coupler diffused light.

In conclusion, the pattern delineation/fabrication techniques have reached a usable stage and considerable work on complex components should develop in the coming years.

What appears now to be a more fundamentally difficult problem concerns the means of obtaining quantitative data concerning the performance of the various elements.

The non-destructive techniques utilized until now have been examination of light diffused out of the guide (by irregularities) and the use of movable prism couplers which couple all of the guided light out of the guide at a given point.

The first technique mentioned is severely limited since the scattered light is produced by localized irregularities and provides practically no information concerning the local optical electric field strength.

The second technique is practically impossible to implement on a truly microscopic scale, and, in any case, yields very little information concerning the transverse distribution of the guided light.

We have recently [22] described a technique which does permit the detailed examination of optical fields propagating in microscopic structures.

A thin film doped with an organic colorant is deposited upon the guide structure that we wish to examine. When the guide is

excited a portion of the guided light is absorbed by the colorant and gives rise to a fluorescent emission. The fluorescent light is not guided, and is therefore, easily observable. Furthermore the fluorescent emission has a different wavelength then the guided light and one can easily discriminate between the fluorescent light and light scattered out of the guide. Since the absorption and the fluorescence are linear processes, the resultant intensity provides a detailed map of the field distribution.

A cross section of the structure used is shown in Fig 10.

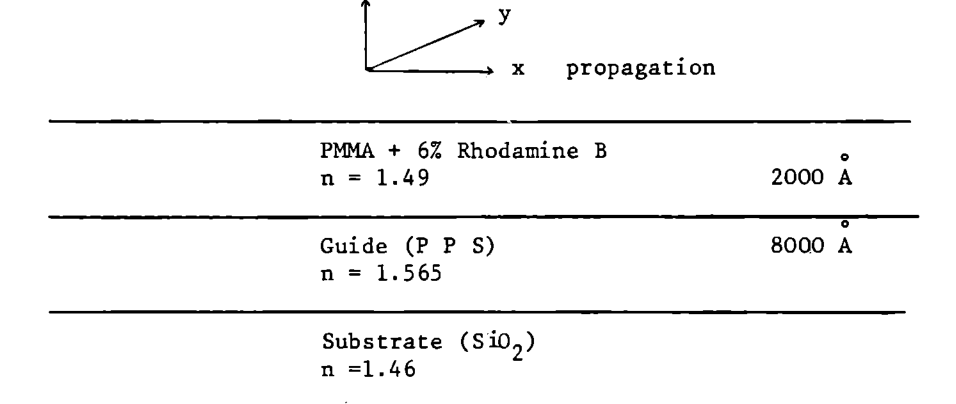


Figure 10 :
Structure for guide examination using locally induced fluorescence

In Fig 11, we see the antistokes (5950 Å) fluorescence emitted where two guides of different widths (20μ and 5μ) have been joined by a triangular transition.

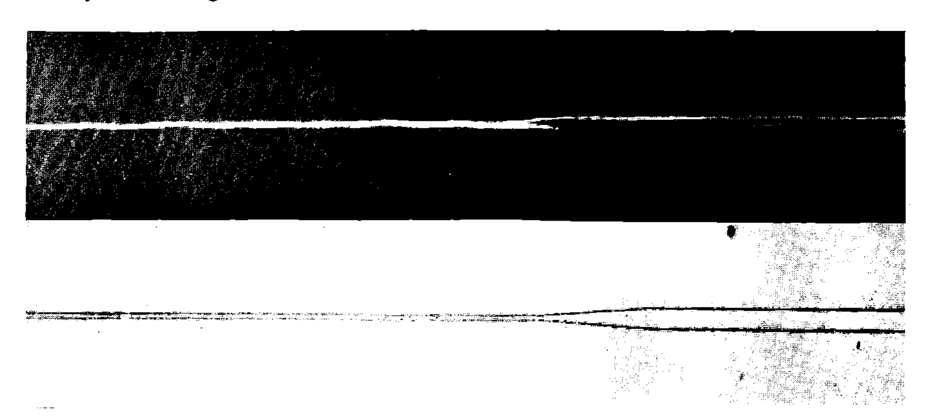


Figure 11 :
Photograph of 20 μ to 5μ transition and fluorescence

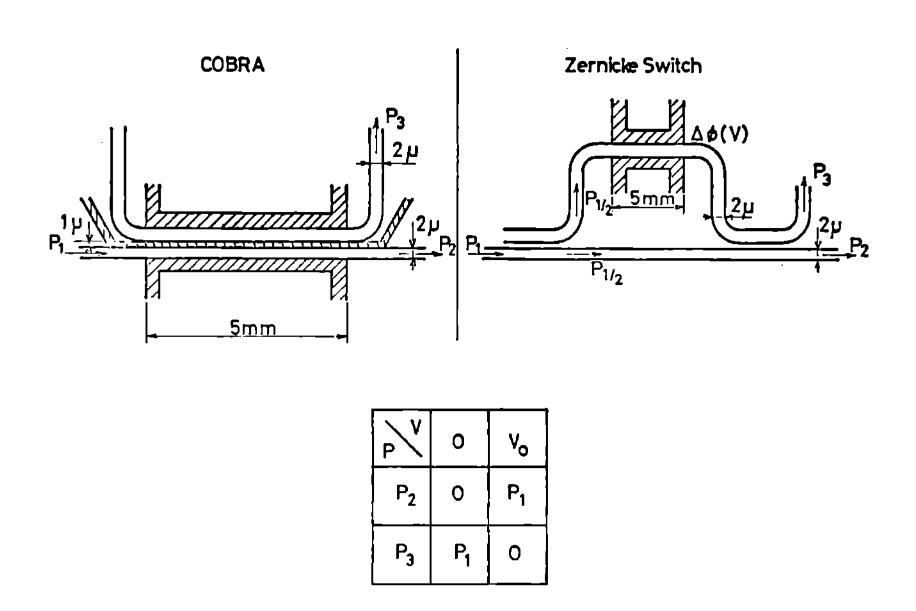
The fluorescence technique permits our visualizing the increased power density obtained in the narrower guide. Quantitative studies of various elements using scanning fiber optic probe measurements of the local fluorescent intensity are now under way and the results of these measurements will aid us in determining the limits of our currently existing delineation techniques, and in the definition of the waveguiding structures for various integrated optical components.

In the following sections we shall discuss in some detail two types of components we can hope to realize using existing integrated optics technology.

▪ V. SOME (DIFFICULT) POSSIBILITIES: THE COBRA AND THE FREQUENCY DOUBLED GaAs LASER

A) The COBRA

COBRA is an acronym, derived from the French and suggested by M. Papuchon (Commutateur Optique Binaire Rapide) which we've adopted for an electronically switchable directional coupler, i.e. a directional coupler realized in an electrooptic material. A schematic of such a device is given in Fig 12.



P \ V	0	V_0
P_2	0	P_1
P_3	P_1	0

Figure 12 :
COBRA and Zernike switch schematics

Due to the evanescent waves extending between them there is an exchange of energy between the two closely spaced guides.

If a power P_1 (given by $|E_1(0)|^2$) is coupled into one of the guides the power in this guide will evolve according to :

$$|E_1(x)|^2 = \frac{|E\ (0)|^2}{\frac{(\Delta\beta)^2}{C^2} + 1} \cos^2\left[\sqrt{\frac{(\Delta\beta)^2}{4C^2} + 1}\ \ Cx\right]$$

where $\Delta\beta$= propagation constant mismatch,

C = coupling coefficient

x = distance along the coupling region

Since the total power is constant the complementary part of the power ($|E_1(0)|^2$ - $|E_1(x)|^2$) has been transferred to the adjacent guide.

Examples of the evolution of energy in guide 1 for two different values of $\Delta\beta$ are given in Fig 13.

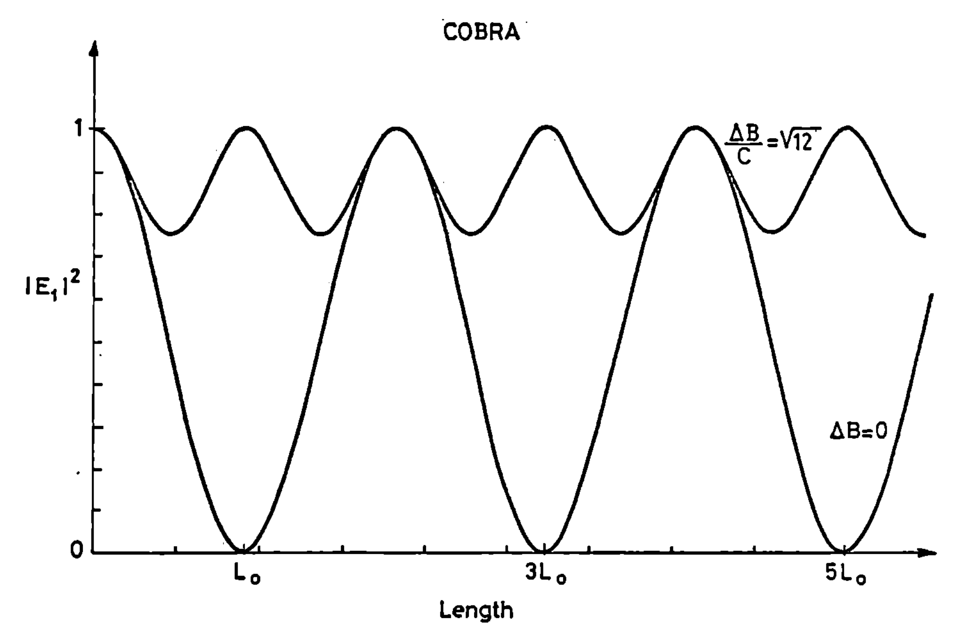


Figure 13:
Evolution of guided energy in one branch of a directional coupler

We see that when $\Delta\beta$=0 all the energy initially in guide 1 will have been transferred to the other guide (at lengths equal to add integer multiples of $L_o = \frac{\pi}{2C}$. When $\Delta\beta$ is non-zero, less

energy is exchanged but the exchance *proceeds more rapidly*.

Evidently we can design a COBRA so that, with a voltage applied all the power leaves by P_3. If by applying a tension we can destroy the resonance, and create a $\Delta\beta$ of approximately $\sqrt{12}$ C all the energy will leave by P_2.

Various ways of implementing this are shown and compared in Fig 14.

Material : $Li\,Nb_x\,Ta_{1-x}\,O_3$

	COBRA	ZERNICKE
Configuration	Command Voltage V	
a) 1 Guide with central electrode	Vo = 1.37 V	Vo = 0.77 V
b) 2 Guides with central electrode	Vo = 0.68 V	
c) 2 Guides with opposed polar	Vo = 1.71 V	
Variation of $\lvert E_2\rvert^2$ with 10 % variation in V	1.6 %	5 %

Figure 14 :
COBRA and Zernike switch comparison

The values of command voltage given are for a COBRA fabricated in $LiNb_xTa_{1-x}O_3$ as suggested by the RCA group [9] .

The actual electrode structures would have to be more complex than those shown, for a and b, to avoid a high resistance central electrode. The structure c is certainly the most attractive but it's not yet clear that the ferroelectric can be polarized in that configuration.

We have also compared the various COBRA configurations with

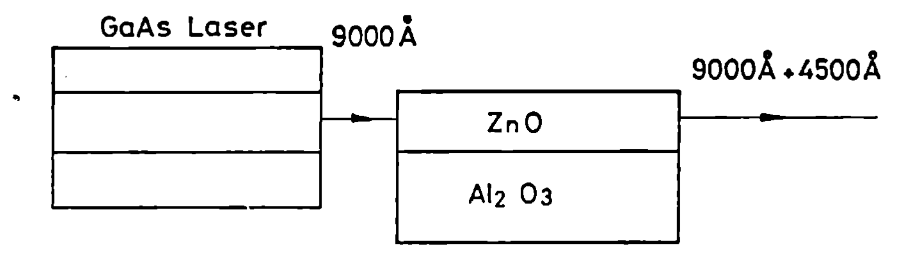
another type of switch suggested by Zernike [23]. Zernicke's switch is essentially an interferometer consisting of two 3 db couplers and a means of changing the phase of the light in one branch. We see that Zernike 's switch will be somewhat more difficult to fabricate (3 db coupling occurs at the steepest slope point of the coupling versus length curve) and more sensitive to control voltage error, but will demand less power than the more practical COBRA's. Evidently the two designs merit investigation.

B) Frequency Doubler for GaAs Laser

In this section we'll discuss an example of a possible "fallout" from the technology developed for integrated optics. By this I mean a device lying outside the mainstream of optical data treatment and transmission systems, which were the original inspiration for integrated optics.

The example chosen is a frequency doubler for a GaAs laser. Such a device could, if it is reasonably efficient, provide an extremely compact source of light in the blue region of the spectrum. Such a source could be useful for reprography, videodiscs, etc.

A simple schematic of such a device is shown in Fig 15.



Frequency Doubler For GaAs Laser

Figure 15 :
Schematic of frequency doubler for GaAs.

In this case we have depicted the frequency doubler (the ZnO/Al_2O_3 guide) as being completely separate from the laser. More highly integrated possibilities exist since it is possible to heteroepitaxy Al_2O_3 and GaAs upon one another.

If we examine the mode dispersion curves for the ZnO/Al_2O_3 guide, Fig 16, we see that for a thickness of about 2600 Å the mode 0 of the fundamental (9000 Å) and the mode 1 of the 2nd harmonic (4500 Å) have the same effective index of refraction ($n \approx 1.80$) and hence, the same phase velocities . This will permit efficient conversion from the fundamental to the second harmonic. The second harmonic power will be given by :

$$P = G[\ \chi\ L\ P_L\ ]^2$$

where :

χ = non linear susceptibility ($\chi_{z\,z\,4z}$)

L = coherence length of system $(L = \frac{\pi}{\Delta\beta})$

P_L = GaAs laser power.

and G is a geometric factor describing the fundamental and second harmonic normalized wave overlap integral. The coherence length will depend on the precision with which we are able to maintain the optogeometric parameters (n , thickness and width) of our guiding structure.

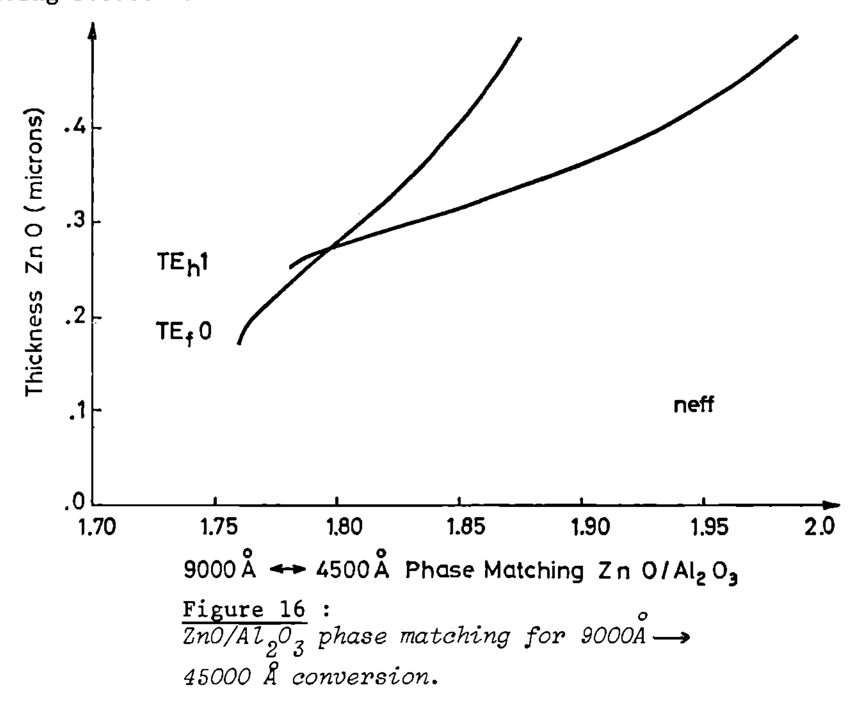


Figure 16 :
ZnO/Al_2O_3 phase matching for 9000Å ⟶ 45000 Å conversion.

It will also depend on the laser spectral width and the related material dispersion. Calculations yield several hundred Å tolerances on the geometrical parameters if the overall structure is coated with a film having an index of about 1.7 . The allowable spectral width can be determined from Fig 17 where we have plotted coherence length versus wavelength. It is clear from this curve that we will want a spectral purity (several Å) better than that one usually obtains with GaAs lasers (10 ∿30 Å).

One possible way to obtain this would be to use a GaAs laser incorporating a distributed feedback cavity as suggested by Kogelnik and Shank [25]. The tolerances on a structure considering

SUBSTRATE X & Indice	Micron	T = 1000 Å		T = 2000 Å		T = 3000 Å		
		Grating period (micron)	T for 1 Å	Grating period (micron)	T for 1 Å	Grating period (micron)	T for 1 Å	Guide Index
X =0.07 3.6435857	0.8	0.10978	3700	0.10977	1600	0.10976	1350	3.651297
X = 0.2 3.5406634	"	0.1126	30	0.1119	26	0.11131	35	"
X =0.29 3.4541702	"	0.11467	9.5	0.11304	12	0.111937	19	"
X =0.07 3.53418	0.9	0.12719	77	0.12688	50	0.12654	55	3.6024071
X =0.2 3.445437	"	0.129917	13	0.128687	15	0.12769	20	"
X =0.29 3.3823658	"	0.131715	8.5	0.129680	10	0.12824	16	"

TABLE II

Tolerances for GaAs Laser on $GaAl_xAs_{I-x}$ *substrate*

the guiding layer in the laser to be GaAs surrounded by symmetric semi-infinite layers of $GaAL_xAs_{1-x}$ are given in Table II.

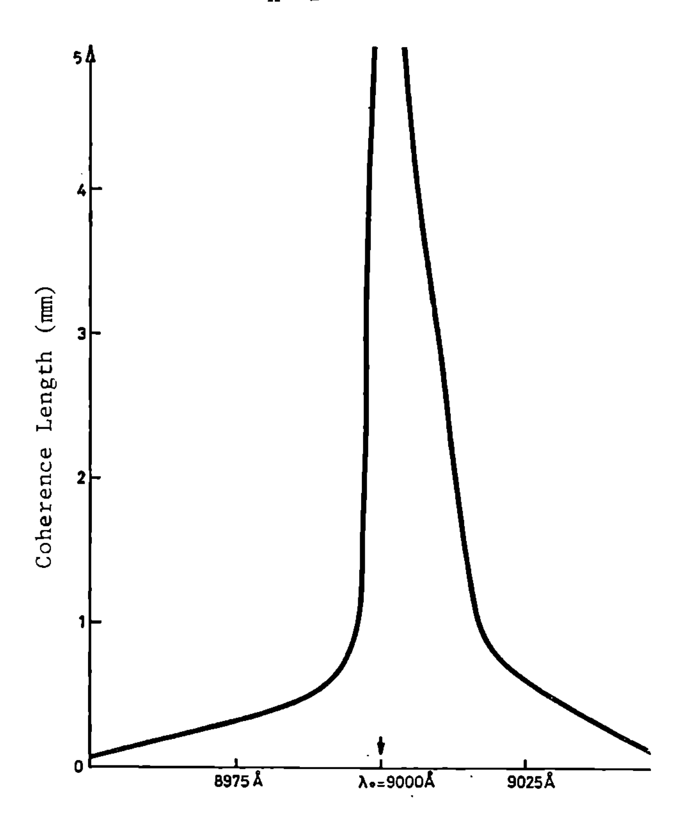


Figure 17 :
Coherence length as a function of GaAs wavelength.

We find tolerances of well under 100 Å on GaAs thickness and *sub Å* tolerances on the grating period. The Δn tolerance is on the order of 10^{-4} .

Despite the obvious difficulties this seems worthwhile to try since some form of tuning (temperature, pressure etc...), should be possible and our calculations indicate that with 10 mW of GaAs power injected into a 10 μ wide ZnO guiding structure , we could obtain 10 ∿20% conversion efficiencies if the system coherence length is on the order of 5 mms.

The challenge is clear.

■ REFERENCES

[1] F.P. Kapron, D.B. Keck, and R.D. Maurer - Appl. Phys. Lett. 17 423 (1970).

[2] I. Hayashi , & al., Appl. Phys. Lett. 17 , 109 (1970)

[3] S.E. Miller, Bell Syst. Tech., 48 2059 (1969)

[4] P.K. Tien, R. Ulrich and R.J. Martin - Appl. Phys. Lett. 14 291 (1969).

[5] D.B. Anderson - Optical and Electro-optical Information Processing (MIT Press, Cambridge, 1965) pp. 221-234.

[6] J.E. Midwinter , I.E.E.E. JQE , 583 (1970)

[7] R. Shubert and J.H. Harris, JOSA 61 154 (1971)

[8] F.K. Reinthart and B.I. Miller, Appl. Phys. Lett. 20, 36 (1972)

[9] J.M. Hammer and W. Phillips - Appl. Phys. Lett. 24 545 (1974)

[10] I.P. Kaminow & al - Appl. Phys. Lett, 24 622 (1974)

[11] P.K. Tien & al, Appl. Phys. Lett. 21 207 (1972)

[12] S.C. Tseng & al, Appl. Phys. Lett. 24 265 (1974)

[13] J.M. Robertson & al, Appl. Phys. 2 159 (1973)

[14] S. Wang, M. Shah and J.D. Crow - I.E.E.E. Trans. Mag. 385 (1971)

[15] S. Yamamoto, T. Makomoto - J.A.P. 45 882 (1974)

[16] D. Marcuse, Bell Syst. Tech. J. 48 , 3187 (1969)

[17] E.A.J. Marcatili , Bell. Syst. Tech. J. 48 , 2071 (1969)

[18] R.V. Pole, S.E. Miller, J.H. Harris and P.K. Tien, Appl. Opt. 11 1675 (1972).

[19] J.C. Dubois, M. Gazard and D.B. Ostrowsky, Opt. Comm. 7 237 (1973)

[20] B. Fay, D.B. Ostrowsky, A.M. Roy and J. Trotel. Opt. Comm. 9 424 (1973).

[21] H. Smith & al.- Elect. Lett. 8, 102 (1972)

[22] D.B. Ostrowsky, A.M. Roy and J. Sevin- Appl. Phys. Lett. 24 ,553 (1974)

[23] F. Zernike - Appl. Phys. Lett. 24 285 (1974)

[24] G. Stillman & al - Appl. Phys. Lett . 24 8 (1974)

[25] H. Kogelnik and C. Shank - Appl. Lett. 18 152 (1971)

CONTRIBUTION A L'ETUDE DES LASERS A STRUCTURE DISTRIBUEE

D.B. Ostrowsky and J. Sevin

Thomson - C.S.F.

B.P. n° 10 91401 ORSAY

I. INTRODUCTION

On pourrait classer les sources de lumière cohérente utilisables pour l'Optique Intégrée en deux catégories : les lasers miniaturisés et les lasers en couches minces. A cette deuxième catégorie appartiennent les lasers " à structure distribuée" (distributed feedback lasers), dans lesquels la cavité optique est obtenue en munissant la couche mince active d'une structure périodique (réseau). Cet article décrit une étude portant sur le fonctionnement de ce dernier type de cavité optique.

Le réseau est obtenu au moyen de variations spatiales périodiques de l'un des paramètres dont dépend la propagation de la lumière. Ce paramètre peut être l'épaisseur de la couche ou du cristal, l'indice ou le gain. Dans les cavités optiques ainsi réalisées, les réseaux jouent le rôle des miroirs des lasers classiques, c'est-à-dire un rôle double de filtre et de réflecteur. En effet, ils servent, d'une part, à éliminer dans le spectre d'émission spontanée les longueurs d'onde qui ne doivent pas laser et, d'autre part, ils contribuent, par réflexion vers l'arrière, à augmenter le parcours dans la cavité des longueurs d'onde à stimuler. Ces deux fonctions de réflecteur et de filtre ne se trouvent pas localisées aux deux extrémités de la cavité comme dans un laser classique, mais au contraire, elles sont réparties tout le long de cette cavité; on dit qu'on a affaire à une "structure distribuée".

Il est probable que ces structures distribuées seront appelées à jouer un rôle important pour la réalisation des sources de lumière en optique intégrée. Il est par suite intéressant d'en comprendre parfaitement le fonctionnement et une série d'essais a été effectuée dans les laboratoires de Corbeville pour les étudier.

Il fallait chercher pour ces essais un milieu actif dans lequel il soit facile de faire varier les paramètres intéressants, tels

que la concentration en dopant et l'épaisseur, et dans lequel il soit possible d'introduire facilement une structure distribuée. Les couches d'essai ont été fabriquées à partir de polyuréthane dopé avec un colorant fluorescent qui est la Rhodamine B.

Ces couches de polyuréthane, réalisées par trempage, présentent plusieurs avantages. Il est facile de faire varier l'épaisseur en faisant varier la vitesse de sortie des plaques du polyuréthane liquide initial. Il est facile de faire varier la concentration en Rhodamine, donc le gain. Enfin il est assez facile d'y imprimer, par pressage à chaud, à partir d'une matrice en Nickel, des réseaux en relief.

L'étude de ces résonateurs à structures distribuées nous permettra d'établir une correspondance entre théorie et expérience, et d'étendre les résultats acquis à l'utilisation de matériaux susceptibles d'applications réelles tels que GaAs et NdUP.

Pour relier théorie et expérience, il faut connaître :
les paramètres de gain du matériau, en l'absence de résonateur,
les caractéristiques spectrales opérationnelles en présence du résonateur = fréquences d'oscillations, largeurs de raies, puissance émise.

Nous avons déjà effectué l'étude des paramètres de gain dans ces couches minces utilisées en amplificateurs (dépourvues de réseaux).

Nous allons, dans la première partie de cet article présenter une théorie et des résultats expérimentaux relatifs aux fréquences d'oscillations et aux largeurs de raies des résonateurs à structure distribuée. Nous ne traiterons pas des mesures de puissance qui sont actuellement en cours dans notre laboratoire.

Dans la deuxième partie, nous discuterons brièvement l'extension de ces résultats au cas d'un laser à résonateur distribué AsGa, à double hétérostructure.

▪ II PRINCIPE GENERAL DE FONCTIONNEMENT DES LASERS A STRUCTURE DISTRIBUEE

A) Etude de la cavité laser

I) Propagation de la lumière de fluorescence dans une couche plastique dopée, sans réseau.

La couche est rendue fluorescente par la présence du colorant, dans une bande de longueurs d'onde assez large (de l'ordre 400 Å). Une partie de cette lumière de fluorescence produite pendant le pompage, se trouve piégée dans la couche par réflexion totale sur les parois et une fraction de cette lumière piégée se trouve émise suivant des angles correspondant aux modes du guide et peut donc se propager.

Cependant, le spectre de fluorescence est continu sur plusieurs centaines d'angströms, et, si la couche est dépourvue de réseau,

la propagation de toute cette bande dans le guide occupe donc, pour un mode donné, une certaine largeur angulaire. D'autre part, le gain résultant de la fluorescence se trouve dans ce cas réparti dans une bande large, et chaque longueur d'onde prise séparément est peu amplifiée (super-radiance).

Les photographies (Fig 1) représentent :

a) projetée sur un écran placé à 50 cm du guide, la lumière de super-radiance sortie de la couche à travers un prisme de recouplage. Les deux raies larges apparaissant sur cette photo correspondent aux deux modes du guide.

b) Le spectre de superradiance continu de 6000 à 6450 Å , photographié à la sortie d'un spectromètre. Les raies de la lampe à Néon, superposées au spectre de fluorescence, permettent la mesure de l'étendue du spectre.

2) Propagation de la lumière de fluorescence dans une couche plastique sur laquelle a été imprimé un réseau.

Le réseau réduit la bande amplifiée, en introduisant dans presque toute la bande de fluorescence des pertes telles que les longueurs d'onde correspondantes ne peuvent plus se propager dans le guide et ne peuvent donc plus concourir à dépeupler les niveaux excités.

Ces pertes résultent de la diffraction vers l'arrière et vers l'extérieur du guide (suivant les ordres du réseau utilisés) de la lumière de fluorescence.

Cependant, pour quelques longueurs d'onde privilégiées et pour certains ordres du réseau, ces pertes sont réduites parce que l'un des angles de diffraction vers l'arrière coïncide précisémment avec l'angle de propagation dans le guide de l'un des modes de cette longueur d'onde. Dans ces conditions, l'énergie qui est perdue dans cet ordre de diffraction pour une onde se propageant de droite à gauche, est récupérée par l'onde symétrique se propageant de gauche à droite, et vice versa.

Le fonctionnement de ce type de laser se trouve schématisé sur le diagramme (Fig 2).

Les deux ondes "D" et "G", de même longueur d'onde, se propagent dans la couche, l'une vers la droite, et l'autre vers la gauche. A mesure que l'une des deux ondes progresse le long de la cavité, elle reçoit de l'énergie de l'onde qui se propage en sens inverse, par diffraction. Il en résulte une réaction (feed back) distribuée tout le long de la structure périodique.

Le gain réparti dans une bande étroite de longueurs d'onde peut être très élevé. A partir d'un certain seuil, il peut donc se produire un phénomène d'oscillation, aboutissant à une émission de raies laser. La photographie Fig 3 représentant le spectre des raies émises est à comparer avec la Fig 1, représentant le spectre continu de superradiance.

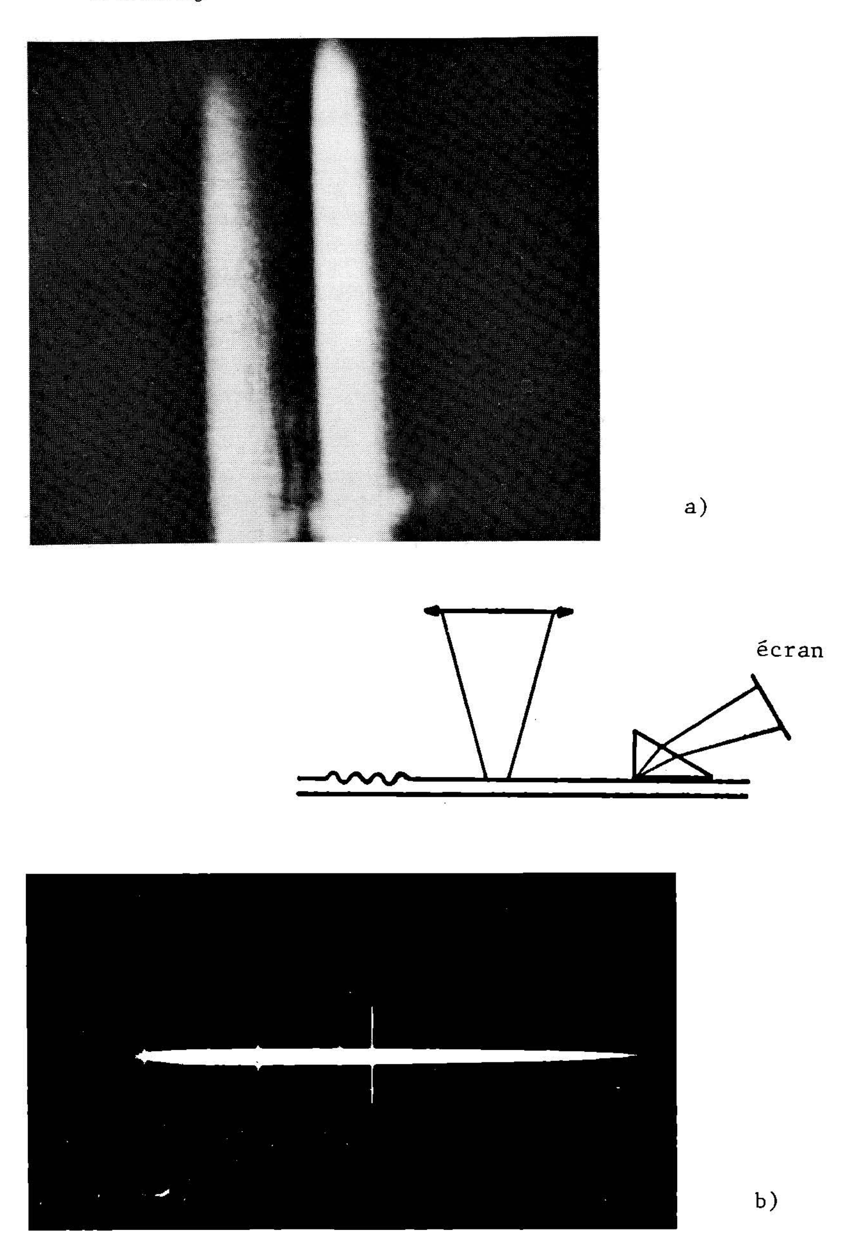


Figure 1
Photographie a): Projection sur un écran de la lumière de superradiance émise au travers d'un prisme.
Photographie b): Spectre de la lumière de superradiance.

LASER A STRUCTURE DISTRIBUÉE

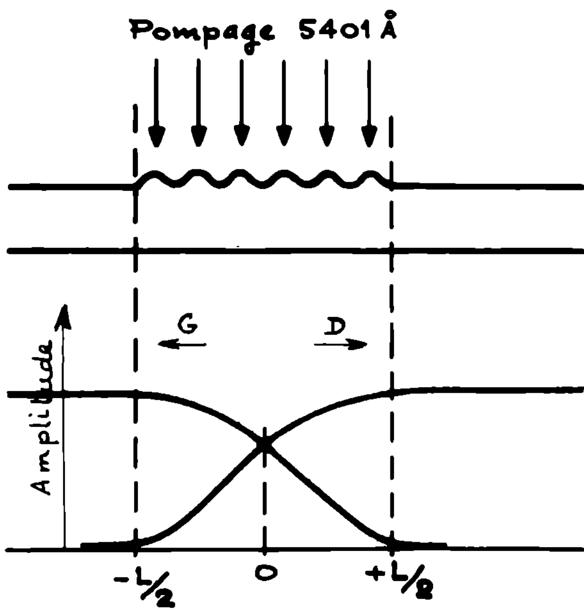


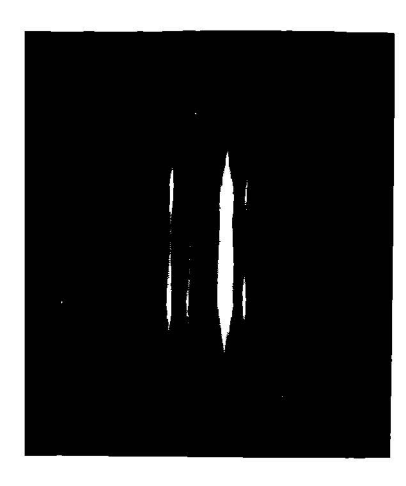
Figure 2 :
Diagramme schématisant le fonctionnement du laser à structure distribuée.

Les mesures de longueurs d'onde nous ont conduits à admettre que les deux ondes contribuant par réaction mutuelle à l'édification du faisceau laser ne se propageaient pas toujours suivant un même mode du guide, mais qu'elles pouvaient dans certains cas, correspondre à deux modes différents du guide. Certaines longueurs d'onde laser résultent donc du feed back d'un mode dans le même mode, et d'autres, résultent du feed back d'un mode dans un autre mode.

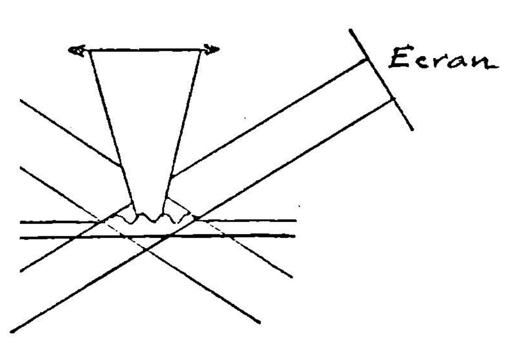
Cette hypothèse a déjà été émise par Zory [4], qui qualifie de longueurs d'onde "intermodes" les longueurs d'onde résultant du "feed back" d'un mode dans un autre, par opposition aux longueurs d'onde "intramodes" résultant du feed back d'un mode dans le même mode.

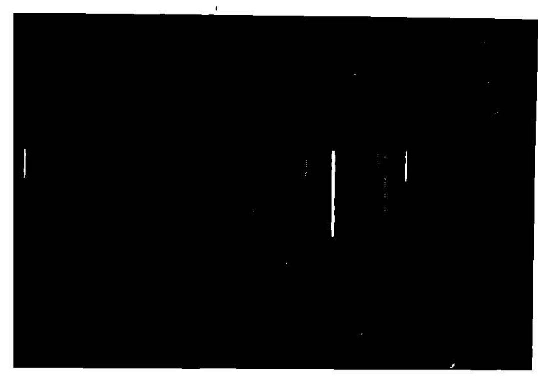
Par raison de symétrie, chacun des deux modes impliqués dans le feed back peut se propager simultanément vers la gauche ou vers la droite. A une longueur d'onde intermode, quatre ondes se propagent donc simultanément dans la cavité, deux vers la droite, et deux vers la gauche, chacune de ces deux directions contenant les deux modes impliqués dans le "feed back" . (Fig 4).

En résumé, une longueur d'onde de fluorescence peut laser lorsque un ou deux des angles discrets de diffraction (correspondant aux différents ordres du réseau) sont égaux à un ou deux angles discrets de propagation de cette longueur d'onde dans le guide (modes).



Emission laser projetée sur un écran placé à 50 cm du réseau





Spectre émis par le laser. Les demi-raies occupant la moitié supérieure du spectre sont celles de la lampe à décharge au néon.

Les raies complètes sont celles du laser

Figure 3 :

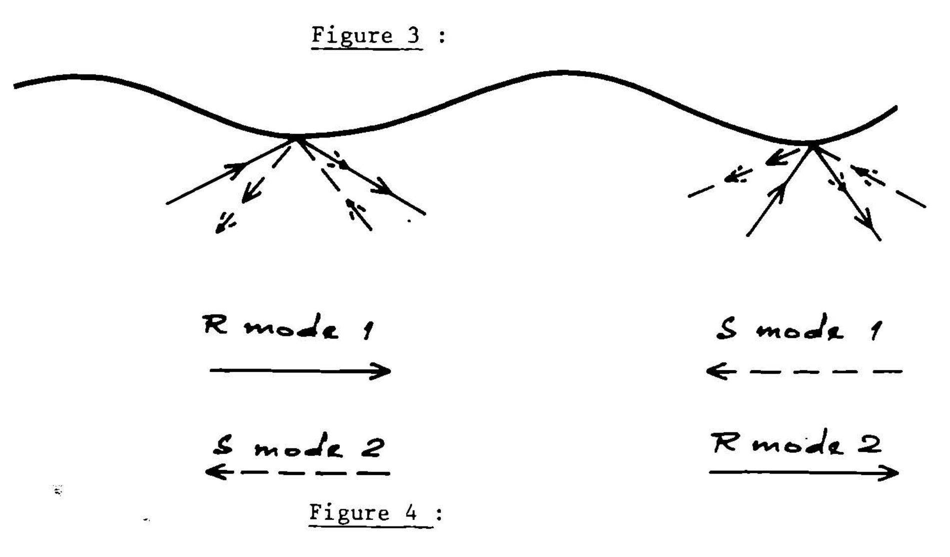


Figure 4 :

B) Vérification expérimentale de ces théories

I) Principe

Les pas des réseaux peuvent être calculés à partir des angles de propagation dans le guide au moyen de formules différentes suivant qu'il s'agit de longueurs d'onde intermodes ou intramodes.

$$2n \sin\theta = \frac{p\lambda}{\Lambda}$$ pour le feed back dans un même mode

et $$n(\sin\theta + \sin\theta') = \frac{p\lambda}{\Lambda}$$ pour le feed back dans un autre mode

et il s'ensuit que les résultats de ces calculs de pas constituent une vérification expérimentale de la validité de ces théories.

La relation entre les angles de diffraction du réseau est donnée par :

$$n_\lambda (\sin\theta + \sin\theta') = \frac{p\lambda}{\Lambda}$$

où n_λ = indice du guide à la longueur d'onde λ

θ = angle correspondant au premier ordre de diffraction

θ' = angle correspondant au deuxième ordre de diffraction

p = entier, ordre de diffraction

λ = longueur d'onde

Λ = pas du réseau

L'équation ci-dessus, donnant θ et θ' en fonction de la longueur d'onde, du pas et de l'ordre du réseau utilisé pour le feed back, se trouve représentée graphiquement ci-dessous.

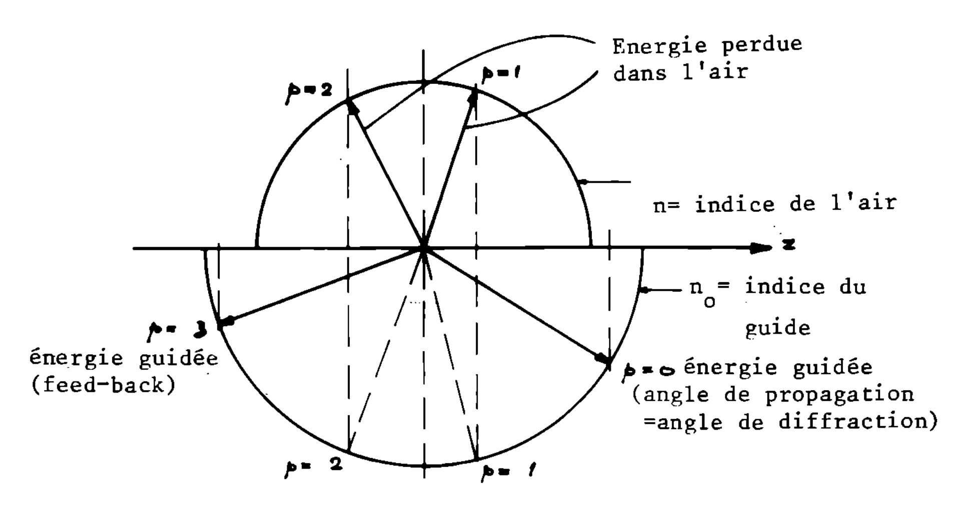


Figure 5 :
Diffraction de la lumière par le réseau dans le cas où l'ordre de feed back =3.

Cette construction graphique donne d'une part les angles de diffraction à l'intérieur du guide, dont deux doivent correspondre à deux modes du guide pour qu'il puisse y avoir feed back d'un mode dans un autre à une longueur d'onde donnée. Si un seul mode est impliqué dans la réaction de feed back, les deux angles doivent être symétriques.

Cette construction graphique donne également les angles d'émergence dans l'air, et montre que ceux-ci correspondent nécessairement à des ordres intermédiaires entre l'ordre 0 correspondant à l'angle de propagation d'un mode et de l'ordre du feed back (dans ce même mode ou dans un autre mode). En effet les angles de propagation dans le guide sont supérieurs à l'angle de réflexion totale à la surface du guide, et ne peuvent donner en aucun cas lieu à émergence dans l'air.

Il s'ensuit que si le pas du réseau est tel que les ordres correspondant aux angles de propagation des modes du guide sont les ordres 0 et 1, il n'y a pas d'ordre intermédiaire, et il n'y a donc pas possibilité d'émergence dans l'air.

En adoptant un pas de réseau suffisamment petit, on peut donc éviter les pertes d'énergie laser par diffraction à l'extérieur du guide.

Pour les longueurs d'ondes intermodes, cette construction graphique ne donne que la moitié des rayons existant effectivement. Par raison de symétrie, il existe une onde se propageant en sens inverse, qui donne par diffraction des angles symétriques.

Pour une longueur d'onde laser résultant de la combinaison de deux modes différents du guide, il existe donc deux angles d'émergence dans l'air vers la droite, et deux angles vers la gauche. (Fig 6).

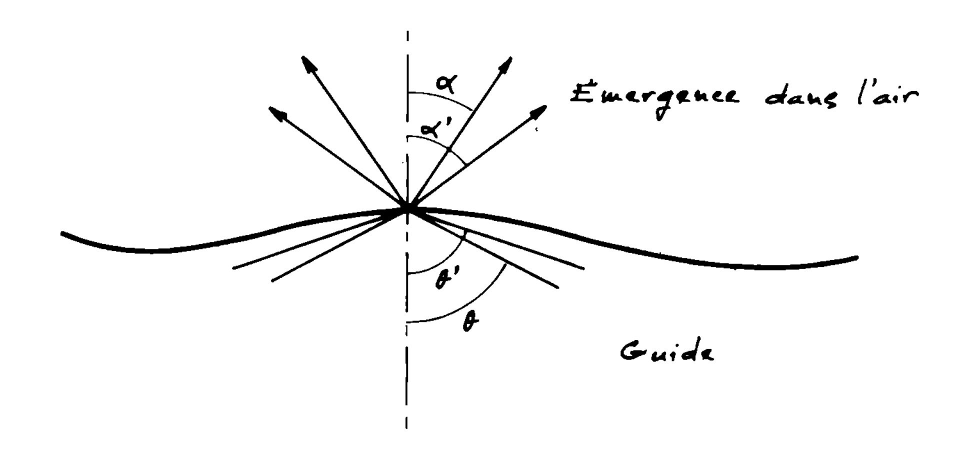


Figure 6 :
Emergence dans l'air d'une longueur d'onde laser résultant de la combinaison de deux modes différents du guide.

Pour une longueur d'onde résultant du feed back dans un même mode, il ne devrait donc y avoir à droite et à gauche qu'un seul angle d'émergence, correspondant à l'angle de propagation dans le guide du mode considéré . (Fig 7).

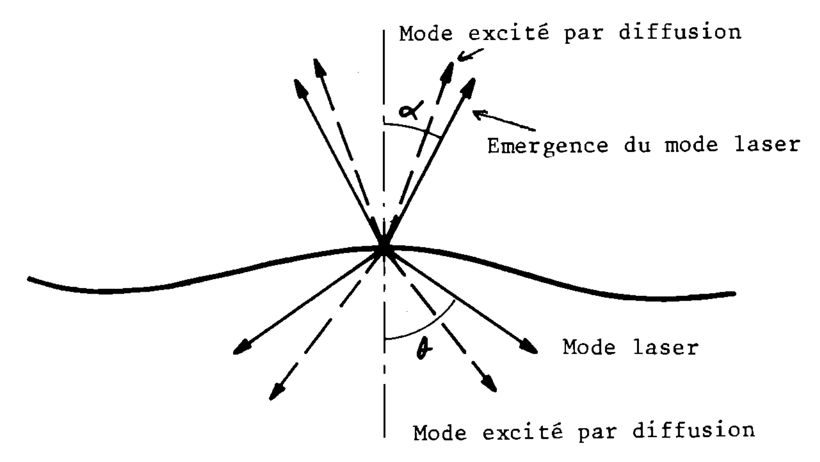


Figure 7 :
Emergence dans l'air d'une longueur d'onde laser résultant du feed back dans un même mode du guide.

Pratiquement, on constate que dans ce cas aussi, il existe deux rayons émergents, correspondant aux deux modes du guide, mais le deuxième mode, qui ne correspond pas à un ordre de diffraction du réseau, est en réalité excité par "diffusion" du premier mode dans le guide.

Cette diffusion résulte des inhomogénéités existant dans la couche

Le deuxième mode est donc introduit par diffusion dans la couche du mode qui lase. Ce deuxième mode ne lase pas, et la différence d'énergie émise dans ce cas entre le premier et le deuxième mode est très élevée.

Dans le cas précédent, au contraire (feed back d'un mode dans un autre), les deux rayons émergents ont la même intensité.

Cette différence dans la répartition des intensités correspondant aux différents rayons émergents est telle qu'il n'est généralement pas possible de confondre une longueur d'onde laser résultant de la combinaison de deux modes, avec une longueur d'onde résultant du feed back d'un mode dans le même mode.

2) Etude du spectre

a) Construction graphique permettant de déterminer les longueurs d'onde laser et leurs angles de propagation :

Pour pouvoir laser, une longueur d'onde soit remplir trois conditions :

1) Elle doit se trouver à l'intérieur de la bande de fluorescence du colorant utilisé pour doper la couche, dans une région de cette bande où le gain est suffisant pour compenser les pertes par absorption et éventuellement diffraction à l'extérieur du guide.

2) Elle doit satisfaire aux conditions de propagation dans le guide d'un ou deux modes. Ces conditions peuvent s'exprimer par une relation liant la longueur d'onde, soit à l'unique angle de propagation, s'il s'agit du feed back dans le même mode ; soit aux deux angles de propagation, s'il s'agit du feed back d'un mode dans un autre.

3) Elle doit satisfaire aux conditions de diffraction du réseau, qui fournissent une deuxième relation entre les angles de propagation et la longueur d'onde.

Les différentes relations existant entre les angles de propagation et la longueur d'onde permettant de tracer des faisceaux de courbes à partir desquels il est possible d'obtenir graphiquement les longueurs d'onde d'émission laser.

Ces différentes relations se trouvent rassemblées dans le tableau suivant où :

w = épaisseur du guide

n_s = indice du substrat

λ = longueur d'onde de propagation

θ et θ'= angles de propagation dans la couche

n_c = indice de la couche

n_a = indice du superstrat (généralement l'air)

m = mode de propagation

<u>Caractéristiques du guide</u>

Pour le mode TM :

$$2\pi W/\lambda = \frac{m\pi + \mathrm{Arc\ tg}\left(\frac{n_c}{n_s}\right)^2 \sqrt{\frac{(n_c \sin\theta)^2 - n_s^2}{n_c^2 - (n_c \sin\theta)^2}} + \mathrm{Arc\ tg}\left(\frac{n_c}{n_a}\right)^2 \sqrt{\frac{(n_c \sin\theta)^2 - n_a^2}{n_c^2 - (n_c \sin\theta)^2}}}{\sqrt{n_c^2 - (n_c \sin\theta)^2}}$$

Les courbes de dispersion caractéristiques pour le mode TE peuvent être obtenues à partir de la même équation en remplaçant les facteurs

$$\left(\frac{n_c}{n_s}\right)^2 \quad \text{et} \quad \left(\frac{n_c}{n_a}\right)^2$$

par 1.

Caractéristiques du réseau

$$2n \sin \theta = \frac{p\lambda}{\Lambda}$$ pour le feed back dans un même mode

et

$$n(\sin \theta + \sin \theta') = \frac{p\lambda}{\Lambda}$$ pour le feed back dans un autre mode.

A partir de ces relations il est possible de tracer des faisceaux de courbes donnant la quantité 2n sin θ en fonction de la longueur d'onde. Les relations caractéristiques du gain donnent un faisceau de courbes, les relations caractéristiques du réseau un faisceau de droites. Les longueurs d'onde pouvant laser sont les abcisses des points d'intersection de ces deux types de caractéristiques.

La figure 8 représente le faisceau de courbes correspondant à une couche de polyuréthane donnée (1,85 micron) de concentration 5,17 gr/l, sur substrat de verre Sovirel B 1664. Le pas du réseau imprimé dans la couche est de 0,62 micron.

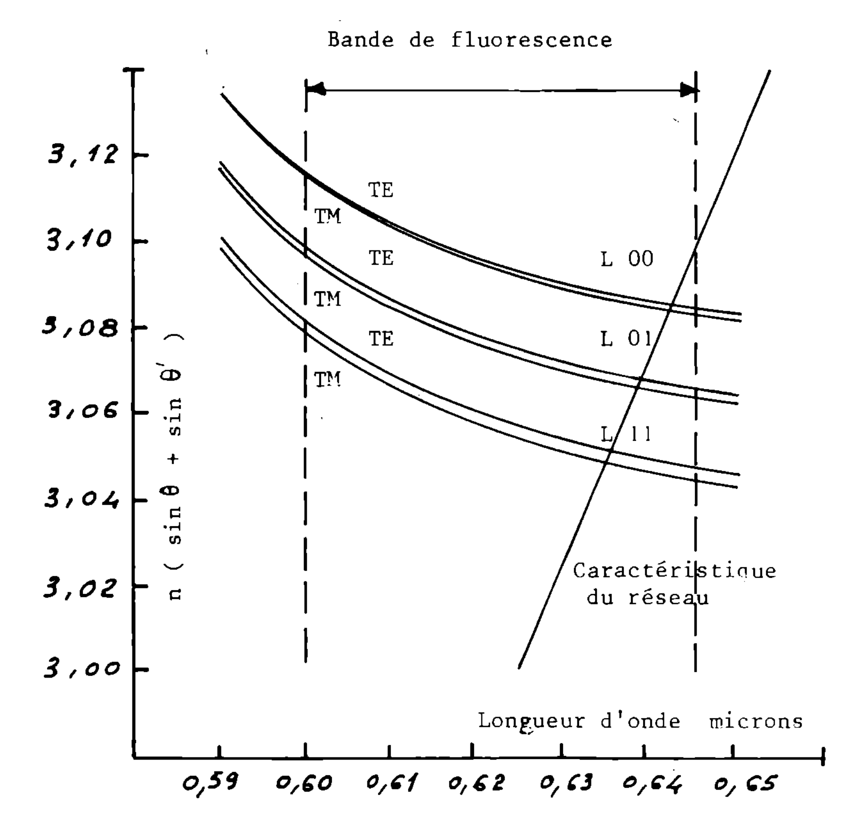


Figure 8 :
Construction graphique des longueurs d'onde laser.

On voit que trois longueurs d'onde peuvent laser en mode TE et 3 en mode TM.

Ces faisceaux sont tracés automatiquement au moyen d'un calculateur Hewlett Packard 9820 A, et du traceur de courbes 9862 A. Les indices de différents substrats et de différentes couches en fonction de la longueur d'onde peuvent être mis en mémoire dans le calculateur au moyen de sous-programmes.

Ces faisceaux de courbes donnent une solution concrète au problème consistant à déterminer les longueurs d'onde et leur angle d'émission.

L'angle de propagation dans le guide d'une longueur d'onde obtenue par réaction d'un mode dans le même mode, est donné sur les courbes par :

$$\theta = \arc \sin \frac{Y}{2n}$$

n étant l'indice à la longueur d'onde considérée,
Y ordonnée du point d'abcisse λ.

Pour une longueur d'onde obtenue par la combinaison de deux modes différents, les deux angles de propagation sont obtenus à partir des ordonnées Y et Y' correspondant à la même longueur d'onde, sur les courbes des deux modes impliqués, (Figure 9) , et sont donnés par :

$$\theta = \arc \sin \frac{Y}{2n} \quad \text{et} \quad \theta' = \arc \sin \frac{Y'}{2n}$$

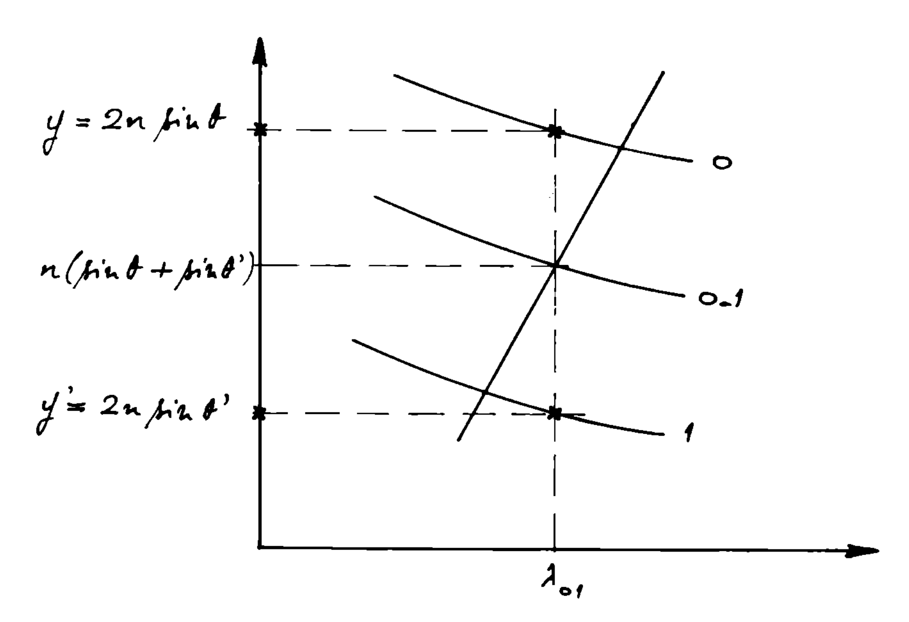


Figure 9 :
Schéma mettant en évidence les directions d'émission d'une longueur d'onde inter-mode sur la construction graphique.

Ces considérations permettent d'expliquer l'aspect des lignes obtenues par projection du faisceau émergent du réseau sur un écran placé à 50 cm . Les angles d'émergence dans l'air croissent avec les angles de propagation dans le guide.

Ainsi, pour un guide permettant la propagation de deux modes ces lignes sont disposées suivant le schéma de la Figure 10.

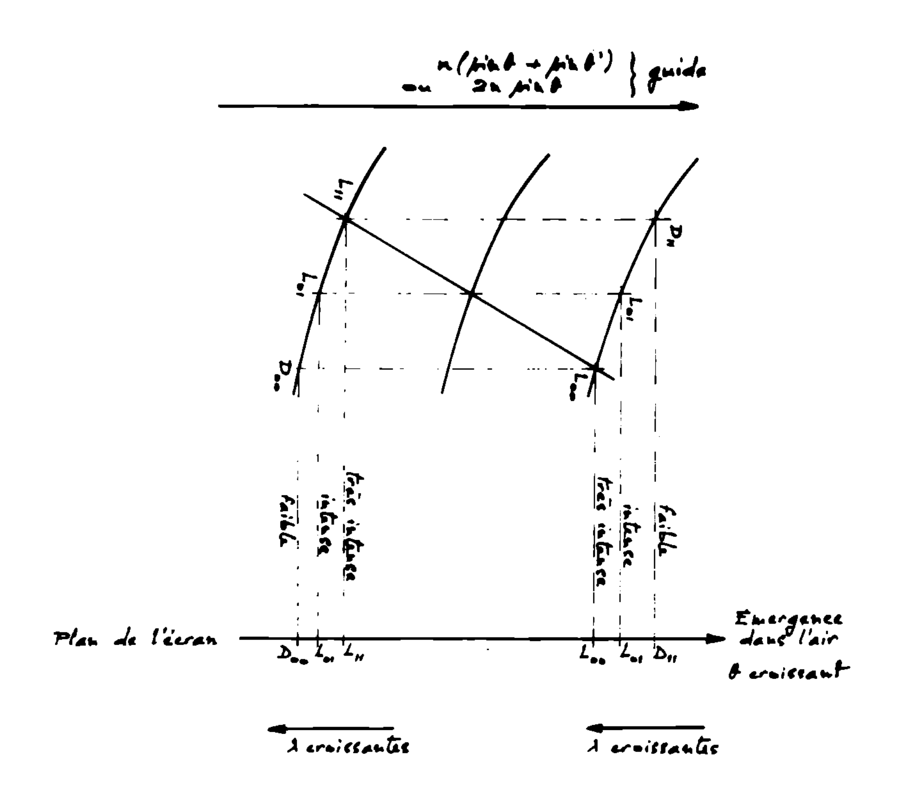
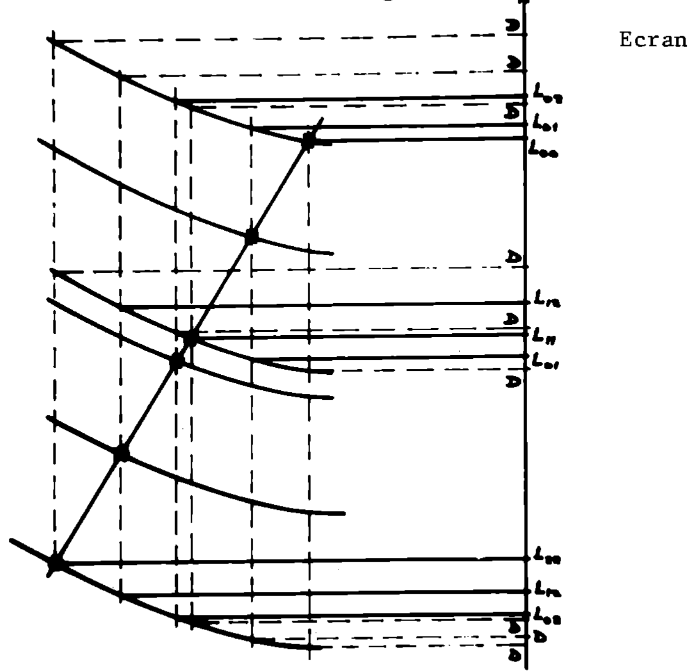


Figure 10 :
Schéma expliquant la disposition des lignes obtenues par projection du faisceau laser sur un écran (guide à deux modes).

Il faut d'ailleurs, pour voir toutes ces lignes que les trois longueurs d'onde laser correspondantes existent effectivement, c'est-à-dire que les points d'intersection qui permettent de les déterminer sur la construction graphique se trouvent tous à l'intérieur de la bande de fluorescence du colorant utilisé, ce qui impose une condition sur le pas du réseau.

Si ce pas est convenable, on observera, pour un guide

permettant trois modes, trois groupes de trois lignes intenses comme on peut le voir sur le schéma de la figure 11.



Projection sur un écran de la lumière émise dans l'air par un laser (guide permettant la propagation de 3 modes).

L= directions d'émission laser

D= directions d'émission de modes excités par diffusion.

Figure 11 :
Schéma expliquant la disposition des lignes obtenues par projection du faisceau laser sur un écran (guide à 3 modes).

Il est facile de trouver quelles sont les lignes qui peuvent subsister sur l'écran, lorsque un ou plusieurs des points d'intersection donnant les longueurs d'onde laser se trouvent en dehors de la bande de fluorescence.

b) Remarques déduites de la construction graphique

La construction graphique page 319 permet de faire certaines remarques :

1) On constate que les longueurs d'onde TE et TM sont très proches l'une de l'autre . La longueur d'onde TM est légèrement inférieure à la longueur d'onde TE, pour un même mode. Cette simple remarque permet de lever les doutes qu'il pourrait y avoir sur l'orientation de l'analyseur utilisé pour identifier les modes TE et TM.

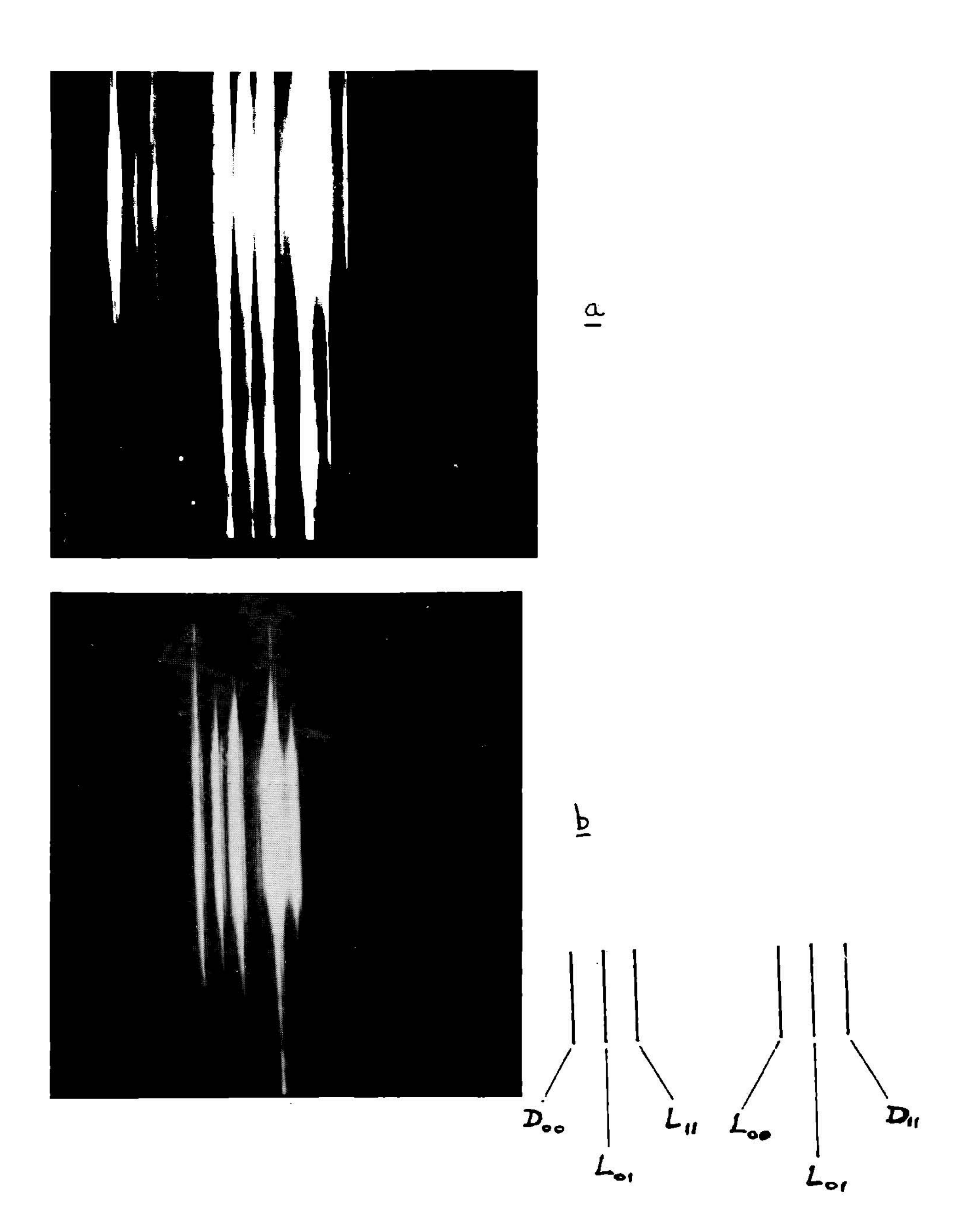


Figure 12 :
Photographies des raies laser projetées sur un écran .
a) - guide à 3 modes
b) - guide à 2 modes

2) Les longueurs d'onde correspondant à la combinaison de deux modes voisins du guide sont à peu près les longueurs d'onde moyennes entre celles correspondant au feed back dans ces deux modes pris séparément.

$$\lambda_{01} \neq \frac{\lambda_{11} + \lambda_{00}}{2}$$

parce que la caractéristique du réseau est une droite, et que, dans l'intervalle de longueur d'onde correspondant à deux modes voisins, les courbes caractéristiques du guide sont aussi assimilables à des droites.

L'écart entre les longueurs d'onde mesurées renseigne donc sur les modes qui leur ont donné naissance.

c) Remarque concernant le nombre de longueurs d'onde émises et le nombre de directions d'émission.

Si le pas du réseau est choisi suffisamment petit pour qu'un seul ordre de réseau participe à la réaction de feed-back, le nombre des longueurs d'onde laser émises par un guide admettant n modes de propagation , est égal à la somme des n premiers nombres entiers :

$$= \frac{n(n+1)}{2}$$

Un guide à 10 modes, peut donc fournir théoriquement 55 longueurs d'onde différentes.

Les modes les moins élevés ont un seuil plus bas,parce que pour ces modes, les pertes résultant des réflexions sur les parois du guide sont plus faibles. Pour faire apparaître toutes les longueurs d'onde prévues par le calcul ci-dessus, il faudrait donc disposer d'une excitation suffisamment forte pour compenser ces pertes pour les modes les plus élevés.

Le nombre de directions d'émission laser (observées dans $\pi/2$ radians) est égal pour un guide admettant n modes à n^2.

Si aux modes laser, on ajoute les modes excités par diffusion, le nombre de directions d'émission possibles devient :

$$n^2 \times \frac{n+1}{2}$$

Un guide admettant 10 modes de propagation émet donc 55 longueurs d'onde différentes dans 100 directions différentes, et si on compte les modes excités par diffusion, ce guide peut émettre les mêmes longueurs d'onde dans 550 directions différentes.

d) Largeur de la bande émise

A une longueur d'onde donnée et à un mode donné, correspond un angle de propagation dans le guide.

Si cet angle s'écarte légèrement ($\Delta\theta$) de l'angle θ de réfraction

vers l'arrière par le réseau de l'un des modes de propagation en sens inverse de la même longueur d'onde, cet écart $\Delta\theta$ fait apparaître des pertes dans la propagation de la longueur d'onde considérée. A chaque longueur d'onde on peut associer un $\Delta\theta$ différent. Lorsque l'onde a parcouru la cavité, elle est d'autant moins amplifiée que le $\Delta\theta$ est plus grand. On aura donc à la sortie du laser un faisceau dont l'intensité varie en fonction de la longueur d'onde (Figure 12 Bis).

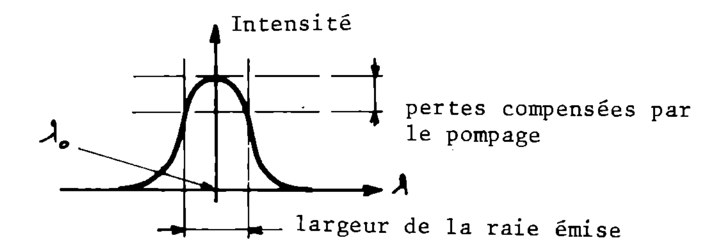


Figure 12 Bis
Rapport entre la largeur des raies et le niveau de pompage.

E) Modes longitudinaux

D'autre part l'écart entre deux modes longitudinaux de la cavité est donné par :

$$\Delta\lambda = \frac{\lambda^2}{2\pi\ell}$$

où ℓ est la longueur de la cavité.

Pour une cavité de 1mm de long (zone pompée) et une longueur d'onde de 0,6 micron, cet écart est donc de 1,2 Å .

Sur la Figure 13 est photographié un groupe de raies correspondant à des modes longitudinaux.

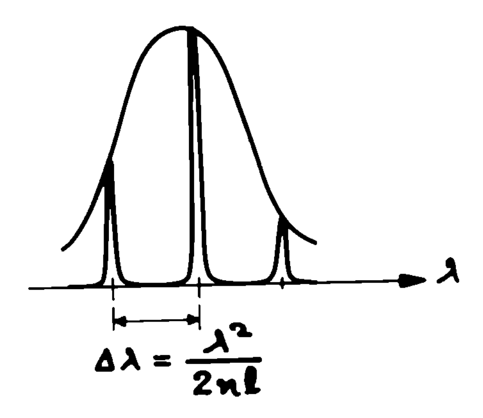
La longueur d'onde centrale est 6140,28 Å , et les deux raies latérales 6137,75 et 6142,76 Å . La longueur de la cavité, correspondant à la longueur pompée au-dessus du seuil pour cette longueur d'onde est donnée par :

$$\frac{\lambda^2}{2\pi\Delta\lambda} = \frac{(0{,}61403 \quad 10^{-3})^2}{2\times 1{,}5548 \times 2{,}5 \quad 10^{-7}} = 0{,}485 \text{ mm}$$

■ III. MESURES D'ANGLE D'EMISSION ET DE SPECTRE

Les mesures consistent à déterminer les longueurs d'onde émises et les angles d'émission correspondants. Les résultats obtenus permettent

de calculer avec une grande précision, le pas du réseau, l'indice de polyuréthane pompé (pour une concentration donnée en rhodamine B) et l'épaisseur de la couche.



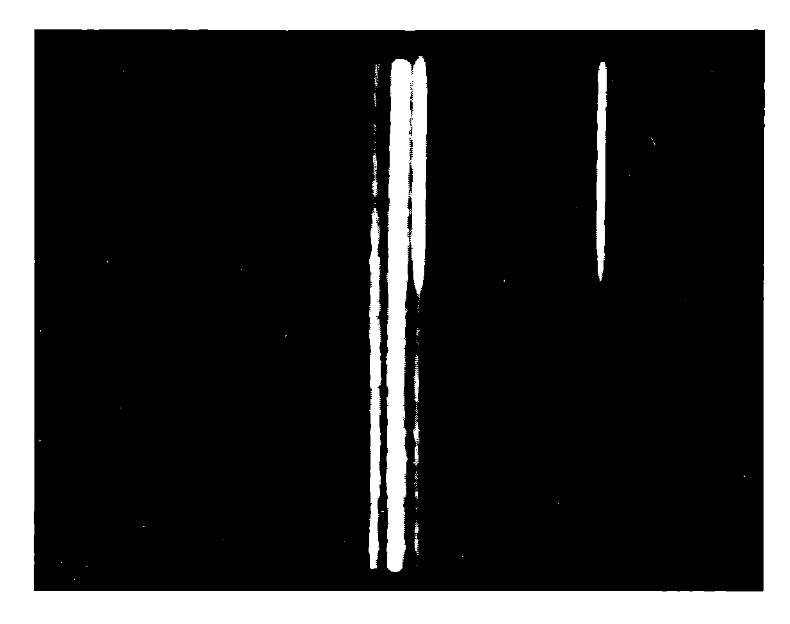


Figure 13 :
Modes longitudinaux

A) Méthode de mesure d'angles

Le montage se trouve schématisé Figure 14.

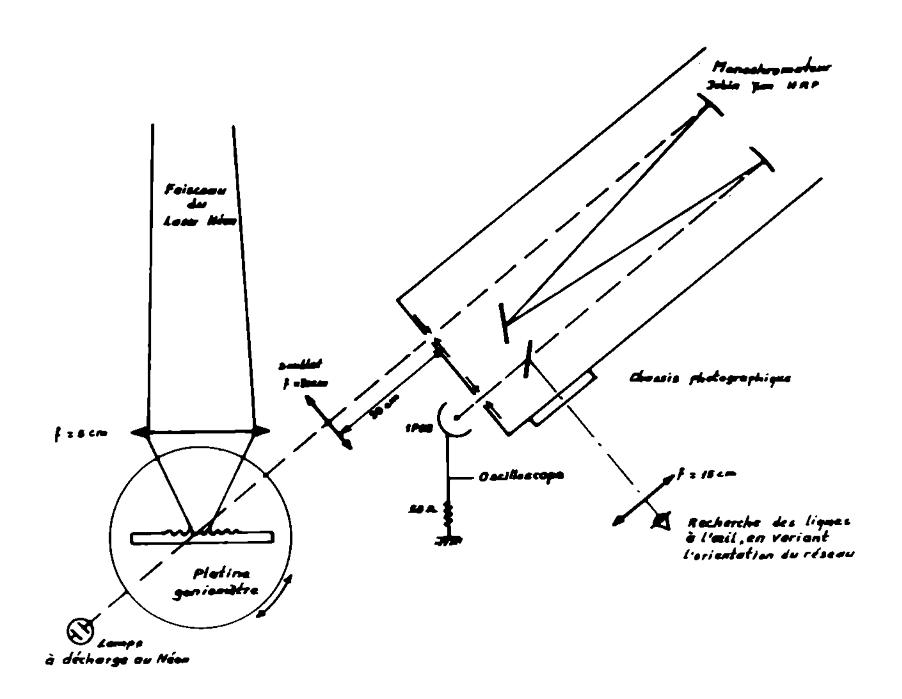


Figure 14 :
Schéma du montage utilisé pour les mesures d'angles d'émission et de spectre.

Le laser en couche mince, monté sur la platine d'un goniomètre BOUTY type 77.000 bis, est pompé au moyen d'un laser au néon AVCO, type C 950 fournissant à 5401 Å , des pulses de 10 KW, de durée 3 nanosecondes.

Pour une longueur d'onde donnée, chaque mode du guide fournit une direction d'émission différente. Le faisceau parallèle émis dans cette direction donne, dans le plan focal du doublet (f =30 cm), une raie verticale, parallèle à la fente d'entrée du spectromètre, et dans le plan de cette fente . Figure 15.

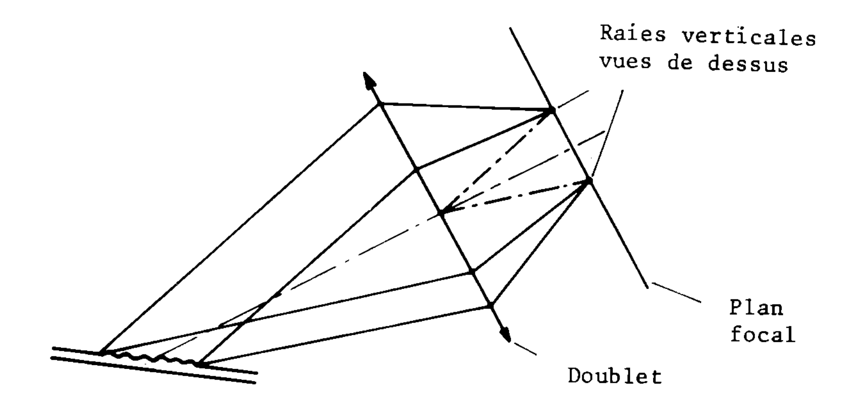


Figure 15 :
Schéma expliquant la formation, dans le plan focal de doublet, de raies correspondant à chaque direction d'émission.

A chacune de ces raies, est donc associée une direction d'émission. En faisant tourner la platine du goniomètre, on fait défiler successivement toutes les raies correspondant aux différentes directions d'émission devant la fente d'entrée du spectromètre.

Le photomultiplicateur RCA , 1P 28 est précédé d'un polariseur ne laissant passer que le mode TE. Le signal observé sur l'écran d'un oscilloscope Tektronix 7904 , est optimisé en procédant à un dernier réglage en longueur d'onde (bouton de réglage du spectromètre), et en angle (platine du goniomètre).

B) Précision des mesures d'angles

Ces mesures consistent à chercher l'orientation de la platine du goniomètre correspondant au maximum d'intensité d'une raie, mesurée au moyen du photomultiplicateur.

Il semble que dans certains cas, la précision de la mesure d'angles soit limitée par la précision du goniomètre (+ ou -20 secondes), et que dans d'autres cas, ce soit le profil de la raie qui introduise la plus grande erreur, le maximum de la raie n'étant pas assez "pointu".

Le profil de la raie peut d'ailleurs se trouver artificiellement aplati pour différentes causes telles que :
- réponse du photomultiplicateur non linéaire,
- lumière diffusée à l'entrée du spectromètre par suite de la présence de poussières, sur les différents éléments optiques du montage.
- raies non focalisées dans le plan de la fente d'entrée, ou non parallèles à cette fente,

- apparition de modes longitudinaux à coté de la raie principale.

Les calculs d'erreurs donnés avec les résultats ont été basés sur une précision de + ou - 30 secondes, qui est pessimiste dans la plupart des cas (les mesures sont généralement moins dispersées que ne le ferait prévoir ce calcul).

C) Mesures de longueurs d'onde, et précision de ces mesures

L'étalonnage en longueurs d'onde du spectromètre prévu par le fabricant ne donne la longueur d'onde qu'avec une précision de + ou - 5 Å .

A partir des photographies du spectre émis, il est possible de faire ces mesures avec une précision meilleure qu'un Å . Le spectre de la lampe à arc dans le Néon fournit des longueurs d'onde repères pour ces mesures.

S'il y a plusieurs modes longitudinaux, le mode central est plus intense, et en réduisant le temps de pose (nombre de coups de laser) et l'ouverture de la fente d'entrée, on peut réussir à repérer avec précision le mode central sur la photographie.

D) Résultats de ces mesures

Les résultats des mesures et des calculs sont donnés pour un laser. Figure 16 se trouvent représentés l'allure du spectre photographié et un schéma situant les longueurs d'onde mesurées sur le faisceau de caractéristiques du laser, et Figure 17 un tableau donne les résultats des mesures et des calculs de pas de réseau, d'indice, et d'épaisseurs.

Ces résultats montrent :
1) Grande précision des calculs
Le pas du réseau est déterminé à 1 Å près, l'indice avec la troisième décimale exacte ; la précision sur l'épaisseur est de l'ordre de 1/10 de micron.
2) Variation d'indice avec la concentration en Rhodamine B
A partir des mesures effectuées sur plusieurs lasers différents, a été tracée la courbe (Figure 18) donnant la variation d'indice mesurée en fonction de la concentration, à la longueur d'onde 6140 Å .
Sur cette courbe on voit que :
n= 1,5539 pour une concentration de 1 gr/l
n = 1,5562 pour une concentration de 3,7 gr/l
En extrapolant, on voit qu'on peut s'attendre à avoir :
n = 1,5565 pour une concentration de 5,2 gr/l
ce qui permet d'avoir l'allure de la courbe donnant l'indice en fonction de la longueur d'onde pour ce pompage et pour cette concentration (5,2 gr/l) , entre 6.000 et 6.400 Å (Figure 19).

L'existence d'une variation d'indice en fonction de la concentration en Rhodamine B nous a conduit à chercher à mesurer la variation d'indice qui pouvait résulter du pompage de ces couches pour une concentration donnée en Rhodamine.

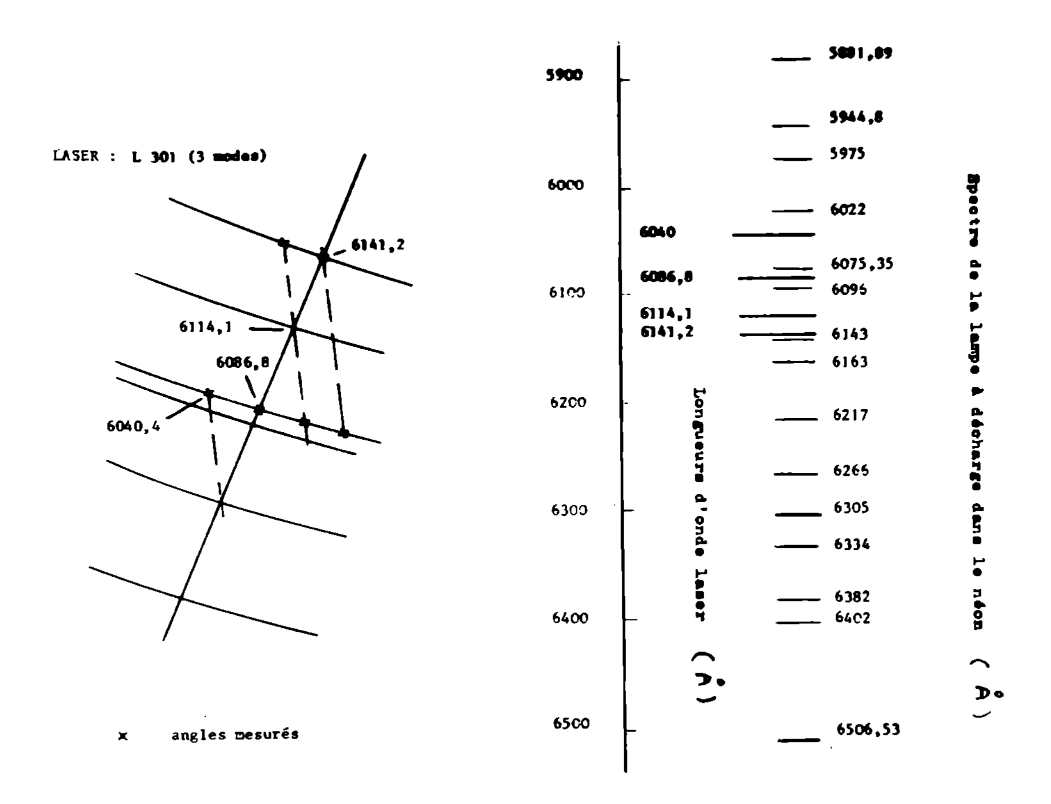


Figure 16 :
Allure du spectre photographié et schéma situant les longueurs d'onde mesurées sur le faisceau des caractéristiques du laser.

La longueur d'onde laser la plus proche du maximum d'absorption (5640 Å) de la Rhodamine, dont nous disposions, est fournie par le laser L 300 (6070 Å), et, par suite de sa concentration relativement élevée (3,3 gr/l), ce laser paraît être le mieux adapté à ce type d'essai.

La puissance du laser de pompage permet de faire varier, pour ce laser et pour cette longueur d'onde, l'énergie de pompage dans un rapport 4 au-dessus du seuil.

Les résultats de ces essais sont négatifs, c'est-à-dire que si le niveau de pompage produit une variation d'indice, cette variation est, dans les conditions de l'expérience, inférieure à la précision des mesures ($0,6 \cdot 10^{-4}$).

Pour réaliser un laser à résonateur distribué, donnant une longueur d'onde précise fixée à l'avance, il faudra tenir compte des variations d'indice résultant de la concentration en Rhodamine.

Resultats des mesures L 30I

Concentration = 1,63 gr/l

Ordre utilisé pour le feedback = 4

λ détectées	Directions détectées	Modes du guide au feedback	Pas du réseau en micron	Indice guide	Epaisseur en micron
0,60404 μ	1 Laser	1 -2			
0,60868	1 Laser	1 -1	0,79222 ±0,00023		
0,61141	0 Laser 1 Laser	0 -1	0,79221 ±0,00022	1,55547 ±0,00077	2,17 0,11
0,61412	0 Laser 1 diffusion	0 -0	0,79229 ±0,00022	1,55521 * ±0,00077	2,12 0,054

*Indice déterminé à partir des angles d'émission laser 00 et de l'angle d'émission de la même longueur d'onde excitée en mode 1 par diffusion.

Noter l'extrême précision des mesures de pas avec ce laser.

Figure 17 :
Tableau donnant les résultats des mesures et des calculs du pas de réseau, d'indice et d'épaisseur.

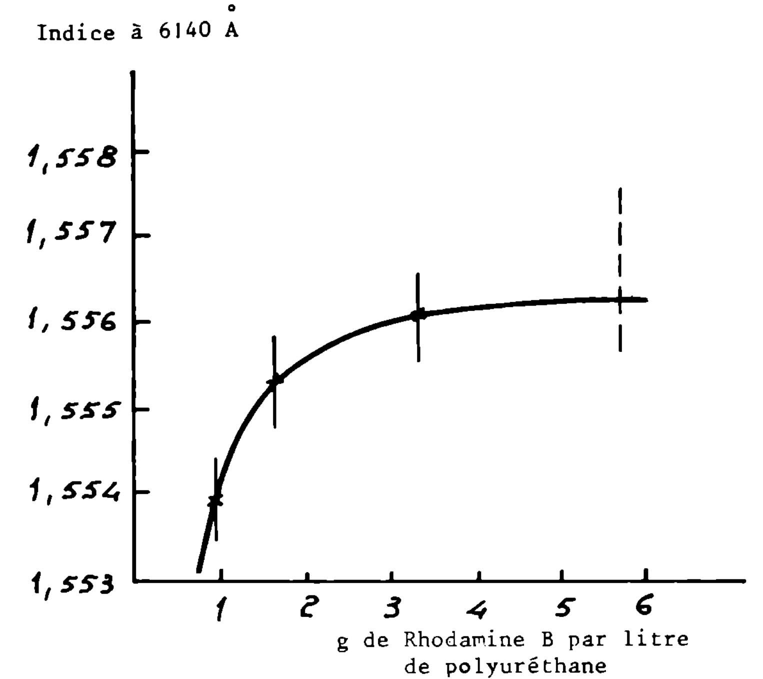


Figure 18 :
Courbe donnant l'indice à 6I40Å du polyuréthane à l'état excité, en fonction de la concentration en Rhodamine B.

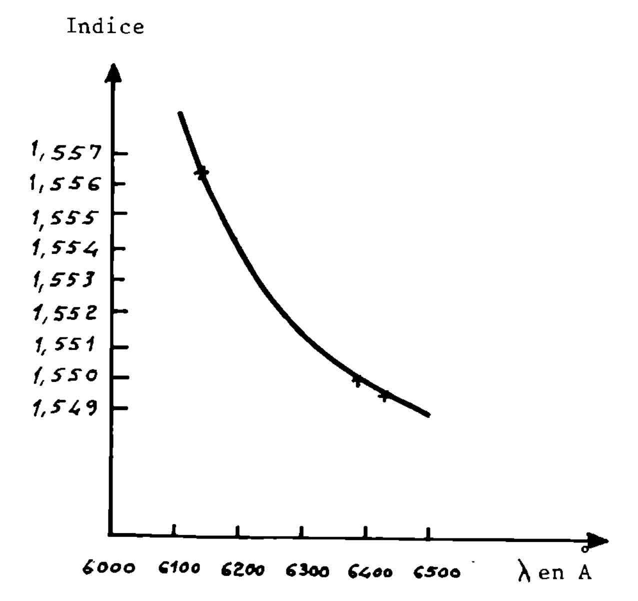


Figure 19 :
Courbe donnant l'indice à l'état excité du polyuréthane dopé à 5,2gr/I de Rhodamine en fonction de la longueur d'onde.

■ IV FABRICATION D'UN LASER EN COUCHE MINCE EMETTANT UNE LONGUEUR D'ONDE FIXEE A L'AVANCE.

Les courbes Figures 20, 21, 22 montrent l'influence sur la longueur d'onde émise des trois paramètres : épaisseur de la couche, indice du polyuréthane à l'état excité et pas du réseau.

Une variation de 10 Å sur la longueur d'onde peut résulter , soit d'une erreur d'épaisseur de 0,1μ , soit d'une variation d'indice de 0,0025 , ou encore d'une erreur de 3Å sur le pas.

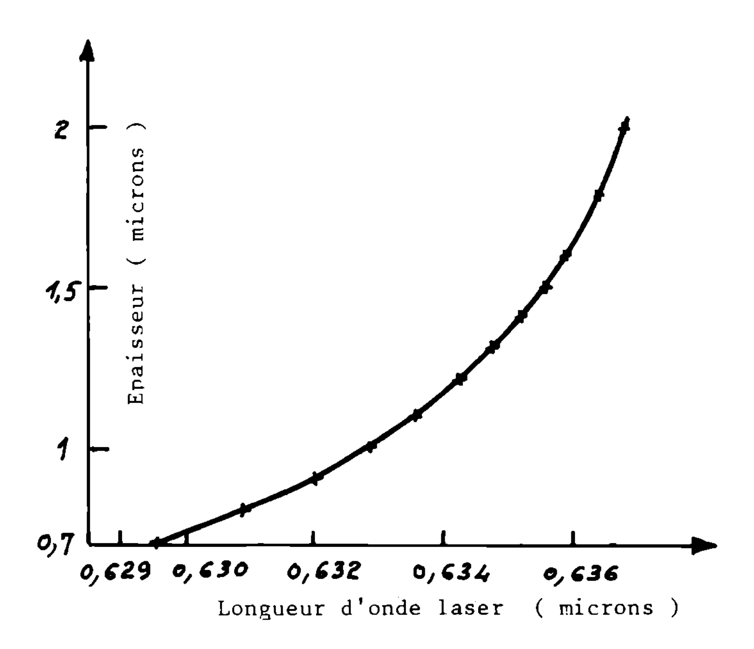


Figure 20 :
Variation de la longueur d'onde émise par le laser en fonction de l'épaisseur de la couche.

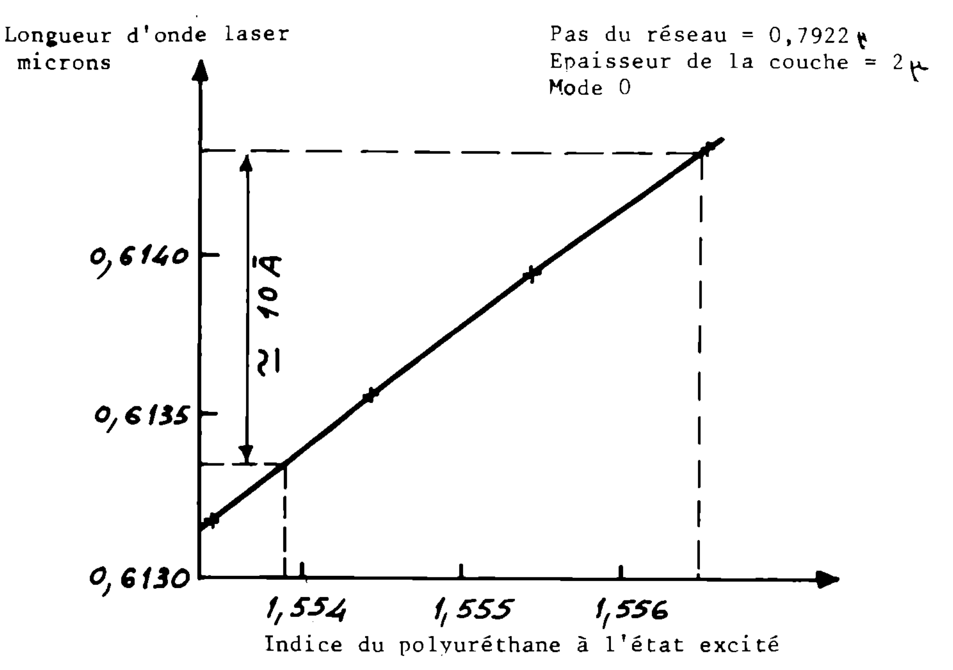


Figure 21 :
Variation de la longueur d'onde émise par le laser en fonction de l'indice de la couche.

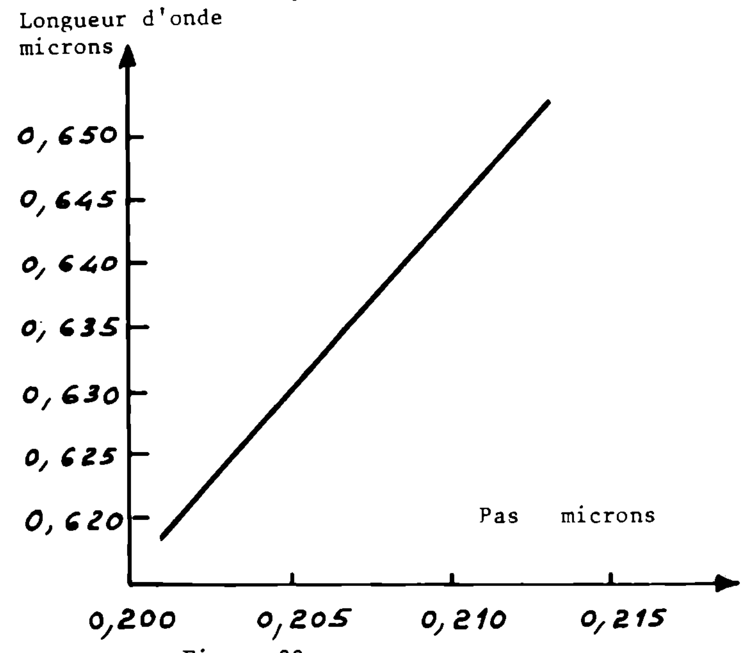


Figure 22 :
Variation de la longueur d'onde émise par le laser en fonction de l'indice de la couche .

▪ V CONTROLE DE LA LONGUEUR D'ONDE EMISE PAR LES LASERS AsGa

La question de savoir s'il est possible d'insérer un résonateur distribué dans un laser AsGa à double hétérostructure est depuis quelque temps à l'ordre du jour. Un tel résonateur devrait permettre de réduire la largeur de la raie, et d'améliorer la puissance émise.

Cependant la réalisation d'une telle structure présente beaucoup de difficultés technologiques, parmi lesquelles on peut citer :
-la fréquence spatiale élevée du réseau (~1100 Å pour obtenir un ordre de feed back = 1) ;
-la nécessité d'introduire le réseau à l'intérieur de la double hétérostructure.

Cependant, Shank et ses collaborateurs [14] ont montré qu'il était possible de réaliser le réseau holographiquement et plusieurs groupes de recherche [15,16] ont montré qu'il est possible de poursuivre la croissance par hétéroépitaxie sur le réseau en relief. Par diffusion [17] il devrait être possible d'introduire la jonction, après réalisation du réseau. Ceci conduit à penser que, en dépit des difficultés, le laser DHS-DFB serait réalisable dans un avenir proche. [15].

Par suite, il vaut la peine d'essayer de voir quelles tolérances permettront de contrôler avec précision, au moyen d'un résonateur à structure distribuée, l'émission spectrale d'un tel laser.

Nous allons considérer ce problème.

Le modèle (idéalisé) est représenté sur la Figure 23. Nous admettons que la région active est constituée d'AsGa pur, placé en sandwich entre deux couches semi-infinies de GaAlAs identiques. Nous avons calculé, pour un écart de 1 Å sur la longueur d'onde émise, les tolérances résultant sur l'épaisseur de l'AsGa, sur le pas du réseau et sur l'indice de réfraction de l'AsGa, dans différents cas. Nous avons considéré des longueurs d'onde laser de 8000 et 9000 Å , des épaisseurs de GaAs de 1000, 2000 et 3000 Å , et des valeurs de X de 0,07, 0,2 et 0,29. Les résultats de ces calculs sont rassemblés dans le tableau 24.

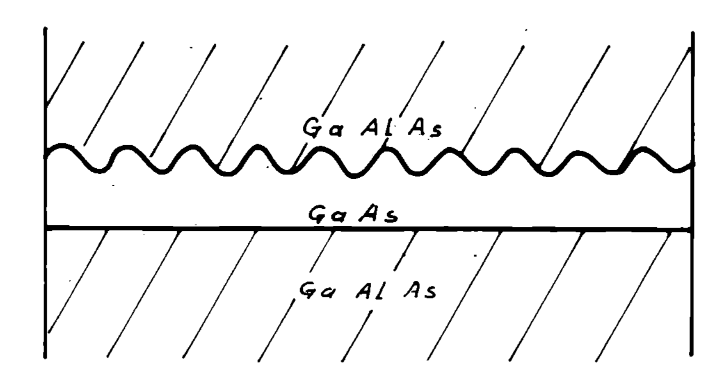


Figure 23 :
Modèle schématique de laser à hétérostructure à résonateur distribué.

λ Micron	Guide indice	Substrat X indice $Ga\,Al_x\,As_{1-x}$	T = 1000 Å			T = 2000 Å			T = 3000 Å		
			Pas réseau micron	Δn	ΔT Å	Pas réseau micron	Δn	ΔT Å	Pas réseau micron	Δn	ΔT Å
0,8	3,651297	X = 0,07 3,6435857	0,10978	0,016	3700	0,109774	0,007	1600	0,1097656	0,0035	1350
		X = 0,2 3,5406634	0,1126	0,0028	30	0,1119	0,001	26	0,11131	0,0007	35
		X = 0,29 3,4541702	0,1147	0,0016	9,5	0,1130	0,00075	12	0,1119	0,0005	19
0,9	3,6024071	X = 0,07 3,53418	0,12719	0,0042	77	0,12688	0,0021	50	0,12654	0,0008	55
		X = 0,2 3,445437	0,12992	0,002	13	0,12869	0,0008	15	0,12769	0,00055	20
		X = 0,29 3,3823658	0,1317	0,0032	8,5	0,1297	0,0015	10	0,1282	0,0007	16

Δn = erreur sur l'indice (
) donnant une erreur de 1 Å sur la longueur d'onde laser
ΔT = erreur sur l'épaisseur (

Tableau 24:
Tableau donnant pour un écart sur la longueur d'onde laser de 1 Å les tolérances sur l'épaisseur et sur l'indice.

Les valeurs calculées dans ce tableau sont théoriques. Pratiquement les indices de l'AsGa pour différents dopages en Al, ne sont pas connus avec une précision supérieure à 10^{-2} . Mais, pour déterminer la longueur d'onde à 1 Å près, il faudrait connaître l'indice avec une précision bien meilleure, comme le montrent les résultats du tableau lui-même. Les calculs ont donc été effectués, sur calculateur, à partir de valeurs d'indices déterminés par extrapolation des courbes de dispersion, en admettant que les indices sont connus avec une précision suffisante pour que la longueur d'onde soit déterminée à mieux qu'un Å près, c'est à dire avec une précision bien supérieure à la précision réellement connue.

De plus, certains des cas présentés ne correspondent pas à une réalité. En particulier, dans le cas où λ = 5 000 Å et X = 0,07, la faible différence d'indice entre guide et substrat ne suffirait pas à contenir le champ électrique assez bien, pour éviter des pertes élevées dans le substrat et le superstrat.

Ce tableau n'a donc pour but que de donner une idée des tolérances nécessaires pour obtenir une précision de ± 1 Å sur la longueur d'onde laser.

Pratiquement, on peut espérer l'obtenir pour :

ΔT = 50 Å

ΔP = 0,1 Å

$\Delta n = 10^{-4}$

La tolérance sur l'épaisseur est dans les limites de ce que permettent les techniques actuelles de croissance par jet moléculaire. Par contre, les deux autres tolérances dépassent nos moyens techniques actuels de contrôle.

En conclusion, on peut dire que dans l'avenir immédiat, l'emploi des résonateurs à structure distribuée pourra être avantageux en ce qui concerne la qualité du résonateur (Q facteur), mais ne permettra pas d'améliorer le contrôle de la longueur d'onde laser.

▪ REFERENCES

[1] D.B. Ostrowsky, A. Jacques, J. Sevin "Etude Optique Intégrée) Marché DRME n° 71-125 Rapport final

[2] M.S. Chang , P. Burlamacchi, C.Hu et J.R. Whinnery "Light amplification in a thin film" Appl. Phys. Lett. Vol 20 n° 8 15 April 1972.

[3] J.E. Bjorkholm, T. Sosnowski et C.V. Shank "Distributed-Feedback lasers in optical waveguides deposited on anisotropic substrates Appl. Phys. Letters Vol 22 n°4 p. 132 Feb. 15 , 1973.

[4] P. Zory "Laser oscillation in leaky corrugated optical waveguides" Appl. Phys. Letters Vol 22 n°4 Feb. 15 , 1973.

[5] J. Bjorkholm et C.V. Shank "Distributed-Feedback lasers in thin-film optical Waveguides" I.E.E.E. Journal of Quantum Electronics Vol QE8 n°11 p.833 Novembre 1972.

[6] D.P. Schinke , R.G. Smith , E.G. Spencer et M.F. Galvin "Thin film distributed-Feedback laser fabricated by ion milling" Appl. Phys. Lett. Vol 21 n°10 p. 494 November 1972.

[7] H. Kogelnik et C.V. Shank "Coupled wave theory of distributed feedback lasers" Journal Appl. Phys. Vol 43 n°5 May 1972.

[8] H. Kogelnik et C.V. Shank "Stimulated emission in a periodic structure" Appl. Phys. Lett. Vol 18 n°4 Febr. 15 1971.

[9] I.P. Kaminow, H.P. Weber et E.A. Chandross "Polymethyl Methacrylate Dye Laser with internal diffraction grating resonator" Appl. Phys. Lett. Vol 18 n°11 June 1st 1971.

[10] H.P. Weber et R. Ulrich "A thin film ring laser" Appl. Phys. Lett. Vol 19 , n°2 July 15 , 1971.

[11] E.P. Ippen et C.V. Shank "Evanescent-field-pumped dye laser" Appl. Phys. Lett. Vol 21 n°7 October 1st 1972.

[12] P. Burlamacchi et R. Pratesi "Waveguide superradiant dye laser" Appl. Phys. Lett Vol 22 n°7 April 1st 1973.

[13] K.O. Hill et A. Watanabe "A distributed-feedback side coupled laser" Optics Communications Vol 5 n°5 August 1972.

[14] C.V. Shank and R.V. Schimdt Appl. Phys. Lett. 23 154 (1973)

[15] C.V. Shank et al. ; J. Ballontyne et al. Topical Conference on Integrated Optics , New Orleans 1974.

[16] M. Nakamura et al; Appl. Phys. Lett. 24 466 (1974)

[17] M.K. Barnoski , R.G. Hansberger and A. Hee Appl. Phys. Lett. 24 627, (1974)

[18] M. Nakamura , A. Yariv , H.W. Yen and S. Somekh , H. Garvin Optically pumped GaAs surface laser with corrugation feedback Appl. Phys. Lett 22 p. 515 1973.

[19] M. Nakamura , H.W. Yen, A. Yariv , E. Garmire, S. Somekh , H. Garvin Laser oscillation in epitaxial GaAs waveguides with corrugation feedback . Appl. Phys. Lett. 23 p.224 1973.

[20] A. Zlenko, A. Prokhorov, V. Sychugov "Thin film layers with distributed amplitude -modulated feed-back"- zh ETF Pis. Red. 18 n°3 156-160 August 5 1973.

[21] C.J. Ernst and W. Witteman - I.E.E.E. Journal of Quantum Electronics Vol QE9 n°9 Sept. 1973 "Mode structure of active resonators"

[22] Stephen R. Chinn "Effects of Mirror Reflectivity in a Distributed-Feedback Laser" I.E.E.E. Journ. of Quantum Electronics Vol QE 9 n°6 June 1973.

[23] A. Yariv and H. Yen Bragg amplification and oscillation in periodic optical media - Optics communications Vol 10 n°2 Febr. 1974 p.120.

[24] S. Chinn and P. Kelley "Analysis of the transmission, reflection, and noise properties of distributed feedback laser amplifiers". Optics Communications Vol 10 n°2 Feb. 1974 p.123

[25] D.R. Scifres, R.D. Burnham and W. Streifer
Appl. Phys. Letters 25 , 203 , 1974 .

OPTICAL TELECOMMUNICATION SYSTEMS USING FIBERS

Teiji UCHIDA
Nippon Electric Co., Ltd
Japan

I. INTRODUCTION

Since the invention of lasers in 1960, the potentiality of optical communications has been discussed for a long time. In the past optical transmission lines, light sources, photo-detectors, and other devices necessary to compose practical communication systems were not developed enough to make an optical communication superior to conventional electrical communication. The recent advances in optical fibers and semiconductor lasers, however, certainly promise the potentiality of optical fiber communications.

Anyone can easily conceive that the usage of optical fibers permits one to construct a reliable optical communication system. But the following classical common-sense in optics had laid this idea aside ; even an excellent quality of optical glass presents an optical attenuation of about 0.1% per centimeter, that is, about 500 dB/km. In 1966, Kao et al. of STL suggested that the elimination of impurities in glass makes it possible to get low loss fibers and that silica presents lower loss performance than glass [1]. In 1968 a light focusing fiber named SELFOC® was realized by Nippon Electric Company and Nippon Sheet Glass Company ; its attenuation was in the order of 100 dB/km for short length at that time [2]. The break-through came in 1970 ; Corning Glass Works produced low loss fibers of about 20 dB/km using silica as mother material. Since then, the intensive research and development of fibers for communication use have been being made in many countries. The latest progress of fibers in Bell Laboratories, and elsewhere, will be introduced by another speaker in this conference.

As to the light sources, semiconductor lasers are probably the best one for communication use ; its size, efficiency, wave-length, and possibility of direct modulation are perfectly suited to fiber communication. The break-through of its degradation problem was reported by Bell Laboratories [3] and Nippon Electric Company [4].

In addition to the above-mentioned technological advances, the incidental coincidence among the wave-lengths of GaAs-GaAlAs double hetero-structure lasers, that of low attenuation area of glass and that of high sensitive area of Si-avalanche photo-diodes has greatly advanced the feasibility of optical fiber communications.

While optical fiber communications will certainly be realized as practical operating system in the near future, it is to be noted that there are still many problems to be resolved such as the long-term equipment reliability in field use, fiber cabling and splicing, installation of fiber cables, maintenance, etc., apart from technical performance improvements.

■ II. FEATURES OF OPTICAL FIBER COMMUNICATIONS

Optical fiber communications have the following remarkable features in comparison with conventional electrical communications.

♦ 1. Broad transmission band-width.
Beside a large modulation band-width at a single carrier wave-length, the band-width due to wave-length multiplexing is extremely large.

♦ 2 . Flat and temperature-variation free frequency response.
This leads to a simple repeater system without equalization.

♦ 3 .Flexibility of fiber itself.
This is particularly important from the view point of economical installation.

♦ 4 . Light weight of fibers.
This favors fiber communications in aircrafts and vertical installation of fiber cables in tall buildings.

♦ 5. Small cross section of fibers.
An extremely large number of information channels can be accommodated in a definite cross section. This may be important in future picture-phone intra-city systems, etc.

♦ 6. Small cross talk.
A high density of fiber bundle cable becomes possible without precaution in designing.

♦ 7 Potentiality of low cost.
The portion of material cost in total cost of fiber cables is very low in contrast with the case of conventional copper cables.

♦ 8 . Non-inductive property.
Fiber communications can be installed near high-voltage or high-current power cables and equipments without special precaution.

♦ 9. Possibility of unipolar connection.
Bipolar connection is not necessary in virtue of optical signals.

♦ 10. Abundant natural resources of glass materials.
Copper resources will be exhausted in the next century.

In addition to the foregoing features, fiber communications have a remarkable potentiality which has never existed in conventional communication systems as follows. Fiber communications will give a broad impact on the fields of almost all kinds of cable communication systems such as on-premise communications, subscriber lines, short and long distance systems, power line carrier sys-

tem, those in aircrafts and ships. This greatly facilitates the problems of manufacturing, installation and maintenance of fiber communication systems, as well as future grade-up of system services, because the same or similar devices regardless of system type will possibly be employed in a wide variety of systems.

▪ III. SYSTEM DESIGN

The mechanism of noise generation in optical receivers is different from that in electrical receivers. Photo-diodes are generally employed as optical detectors. In this case the number of incident signal photons onto a photo-detector is converted into that of output electrons from the photo-detector. This means that a photo-detector is an ideal square law detector. On the other hand, there exists quantum noise due to corpuscular nature of signal photon flow in addition to conventional thermal noise. This quantum noise is changed into shot noise in the output side of the photo-detector. In order to improve the signal to noise ratio in an optical receiver, it is necessary to achieve an internal multiplication of photo-electrons in photo-detector itself, so that an avalanche photo-diode is preferably employed.

Taking account of these facts, an example of system design of an optical fiber communication system using a semiconductor laser and an avalanche photo-diode is given in Fig. 1. In the cases of 100 M bit/s and 1G bit/s transmission capacities, the allowable span losses are about 48 dB and 40 dB, respectively. When a fiber loss is assumed to be 10 dB/km and a pulse-broadening due to fiber dispersion is neglected, the span is on the order of several kilometers. Figure 2 indicates a relationship of span versus information bit rate. A span in which loss limit becomes equal to dispersion limit is preferably chosen as the optimum design.

It should be noted that this span will be reduced in reality by various kinds of system margins.

▪ IV. DEVICES AND GUIDES

Since reviews on each device and guide of optical fiber communications are given in detail by other authors, the description in this chapter is mainly concerned with the works of the author's group.*)

A. Transmitter

1) *Choice of optical sources*

In the present stage, three kinds of optical sources are generally considered to be suitable for communication use : semicon-

---

*)Readers, who want general reviews on optical fiber communications are also refered to the paper by Miller et al [5].

ductor lasers, light emitting diode (LED), and LED pumped YAG lasers.

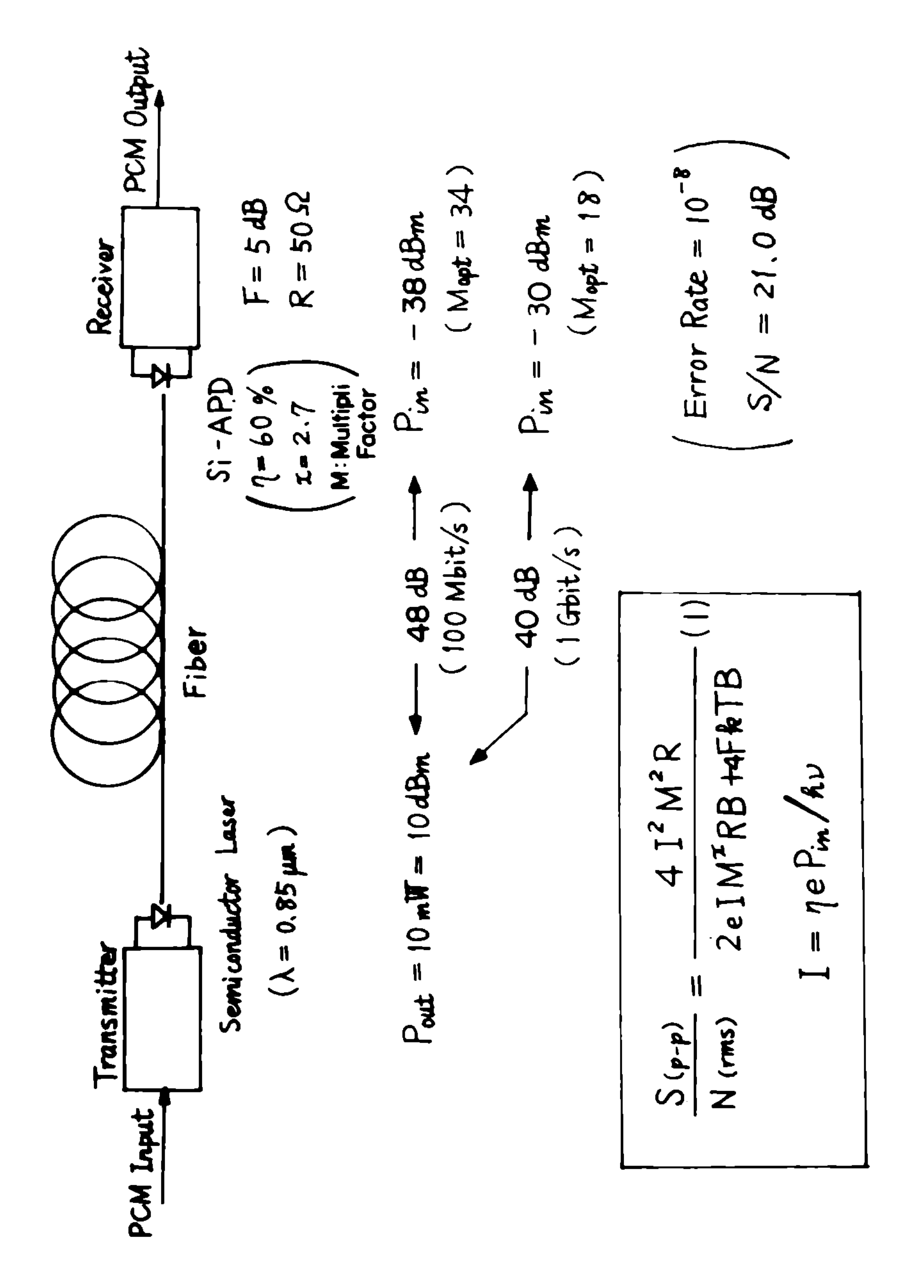


Figure 1 :
An example of system design of an optical communication system.

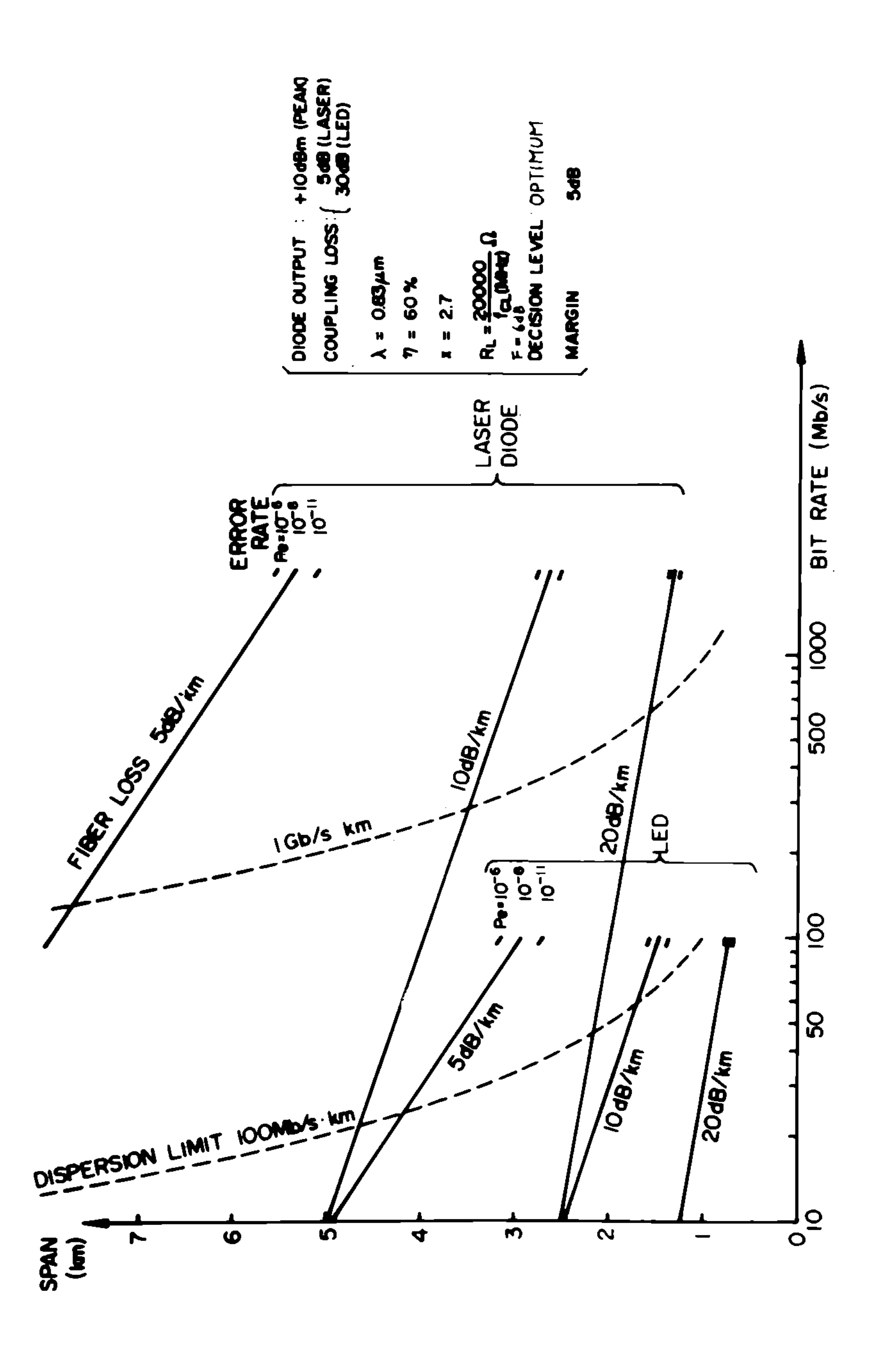


Figure 2 :
Span vs. information bit rate.

The Nd doped YAG laser pumped by LED [6] gives a wave-length of 1.06 μ at which glass presents lower loss than that at 0.8 microns of semiconductor laser emission. Another feature of this laser is to yield supposedly a long life on account of LED pumping. It, however , gives rise to the following disadvantages : need of an external optical modulator, high cost, and low efficiency. By these reasons, its applicability to optical fiber communication systems is not clear at present.

Table I indicates the comparison of DH (double hetero-structure) laser and LED for communication system. The low coupling efficiency of LED is particularly serious from the viewpoint of system design. Furthermore, the differences of cost and life between DH laser and high brightness LED specially designed for communication use are likely to disappear in the future mass production stage.

	DH LASER	L E D
SPECTRAL WIDTH Δλ	∿20 Å	∿300 Å
MODULATION BANDWIDTH	∿1 Gbit/s	∿100 Mbit/s
COUPLING LOSS TO FIBER	< 3 dB	∿17 dB

Table I :
Comparison of DH (double hetero-structure) laser and led

The degradation problem of DH lasers is being resolved rapidly in Bell Laboratories ,[3] Nippon Electric Company, [4] and elsewhere. It will be expected to achieve the life longer than tens of thousands hours in future.

2) *Coupling between a DH laser and an optical fiber.*

In general, the number of simultaneously propagating transverse modes through a fiber must be small to minimize the pulse broadening due to mode-dependent group delay, so that an excitation of the fundamental mode of a fiber is ultimately preferable. This fact leads to the single transverse mode operation of a DH laser

using a stripe-geometry structure.*) Since the stripe-width of this structure is considerably large compared with the thickness of the active layer of the DH laser, the near-field pattern of an output laser beam is extremely elliptical. Accordingly a special precaution must be paid to the coupling between a DH laser and a fiber. Figure 3 shows an example of coupling configuration for this purpose .[7] The first cylindrical slab SELFOC lens converts a long-elliptical output mode of a DH laser to a circular

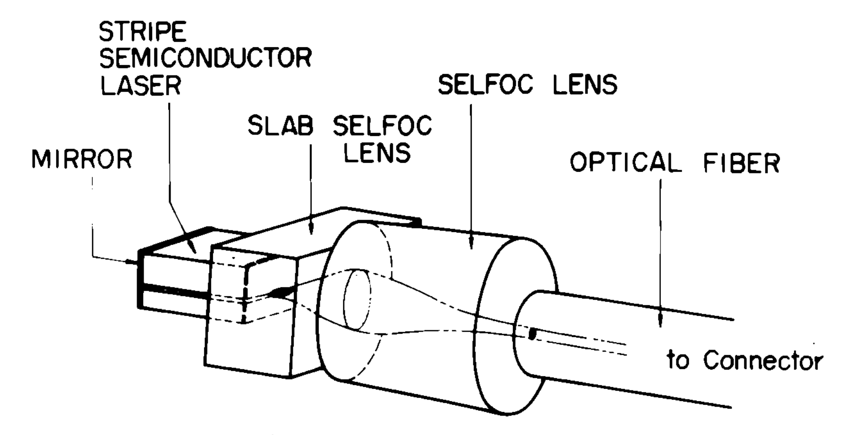


Figure 3 :
Mode matching between a semiconductor laser and an optical fiber.

one, which is further transformed to the proper fundamental mode by the second circular SELFOC lens. Recently another configuration was developed to obtain better coupling efficiency. [8] Figure 4 shows this configuration using two perpendicularly crossed cylindrical slab SELFOC lenses. Compared with the foregoing one, this configuration renders the better tolerance as to lens positioning, the wider covering range against the deviation of radiation angle of DH lasers, etc. The measured coupling loss is given in Table 2 for DH lasers with different radiation angles in a plane vertical to the junction. Since these coupling losses involve the reflection loss at each end surface of the lenses and the fiber, the coupling efficiency will be ameliorated by about one decibel by anti-reflection coatings. Besides, a portion of the indicated losses are also due to the difference of apertures of the incident laser beam and the first slab lens. Therefore, the coupling loss will be reduced further. Apart from performance, the coupling configuration using SELFOC lenses with flat end surfaces is convenient for practical manufacturing.

---

*) Such an operation is also favorable to the elimination of oscillation instability.

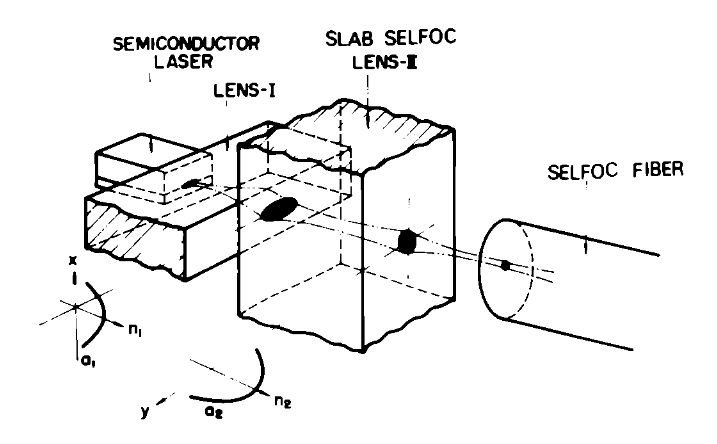


Figure 4 :
Coupling using two vertically crossed slab Selfoc lenses.

RADIATION ANGLE	SLAB SELFOC® LENSES	CIRCULAR SELFOC® LENS
I6 °	2.0 ∿3.5 dB	3.0 ∿4.5 dB
35 °	3.4 ∿4.0 dB	6.2 ∿7.5 dB

Table II :
Coupling loss between DH laser and SELFOC® fiber (include reflection loss)

3) *Modulation of a DH laser*

An important property of DH lasers is the ability of direct modulation through pumping current. Figure 5 shows a typical relationship of light output power versus drive current in a DH laser. While the light output response to modulating drive current in poorly population inverted regime is not fast, it becomes fast in highly population inverted regime near and above threshold. A dc-bias current is, therefore superposed on a modulating drive current. In case of insufficient dc-bias current, a pattern effect is observed. That is the peak amplitude of a laser pulse corresponding to a drive current of signal 1 which follows a sequence of signal zero becomes small compared with those of successive laser pulses. But a dc-bias should not be too high, because the increase of the dc-bias causes the increase of dissipated power of a DH laser

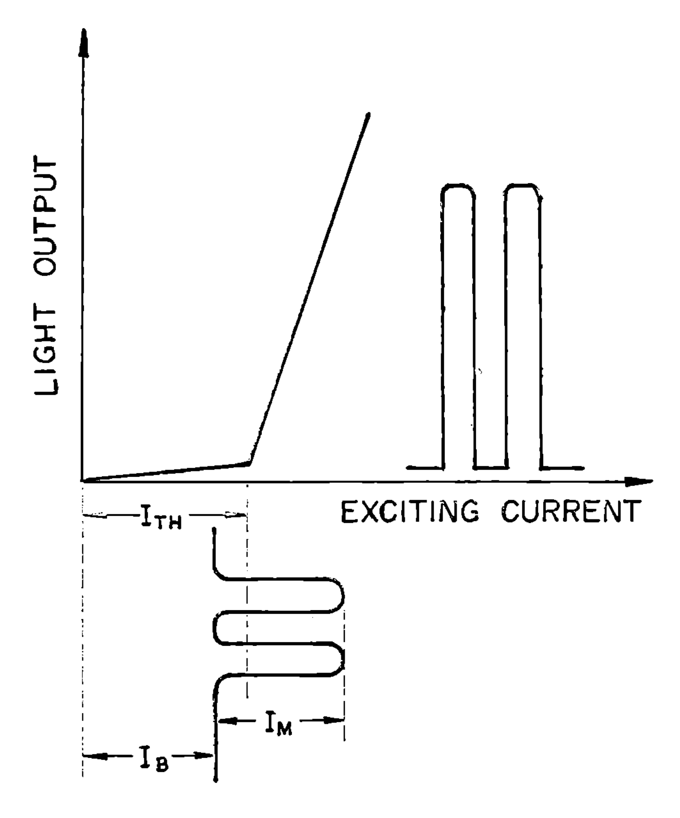


Figure 5 :
Current modulation of injection laser.

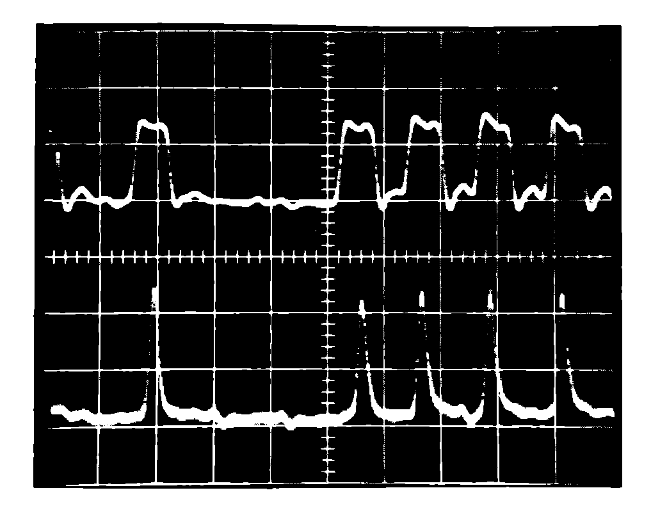
provoking a degradation of the laser.

Such a dc-prebiasing can be effective in reducing a relaxation oscillation of DH lasers. The resonance phenomenon attributed to an interaction between optical field and injected carrier in a DH laser causes a ringing of output laser signals. Figure 6 shows a result of 400 Mbit/s direct modulation of a DH laser properly pre-biased. A typical relaxation oscillation is observed in Figure 7 which indicates the data of 1 Gbit/s modulation. Since the ringing frequency is on the order of GHz, the usage of electrical filtering makes this ringing effect harmless in case of lower frequency modulation. However, such a relaxation oscillation must be removed for higher frequency modulation.

4) *Stabilization of a laser output power.*

A reliable optical transmitter for operating system requires an automatic level control circuit (ALC) for a DH laser, because the output power of a DH laser is susceptible to temperature variation and the laser gradually deteriorates with time. For this purpose, a laser beam emitted from the opposite end of a DH laser with respect to its output end facing a fiber is detected with a photo-diode, whose current is simultaneously employed for output

power monitoring and feed-back to control the drive current of the DH laser.



Vertical

upper : electrical rz signal

20 mA/div
lower : laser output signal

2 mW/div

Horizontal

2 ns/div

Figure 6 :
400 Mb/s direct modulation of GaAs injection laser.



1 Gb/s MODULATION

Vertical

upper : modulation signal
lower : laser output signal

Horizontal

2 ns/div

Figure 7 :
Typical relaxation oscillation of a DH laser.

B. Receiver

As mentioned in Part III, an avalanche photo-diode is desirably used as photo-detector. Although a p-i-n photo-diode is employed for an optical analog signal transmission on account of its linearite, it loses the amelioration of S/N attributed to internal gain

1) *Avalanche photo-diode (APD)*

The avalanche photo-diode, already reported elsewhere,[9] was designed optimally for use at 6328 Å. While this P^+PNN^+ APD acts as electron-injection type at 6328 Å, the APD becomes hole-injection type for longer wave-length such as 0.8 μ and 1.06 μ. As the consequence, the performances of the APD regarding noise and response become inferior in these infra-red regions.

On the other hand, an APD made by ion-implantation is suitable to the usage at the wave-length of DH lasers. [10] Figure 8 shows the profile of this APD in which the layer N^+ of 0.6μ thickness was made by implanting phosphorus ion into a Si epitaxial wafer of P on P^+. The spectral response of this APD is given in Figure 9.

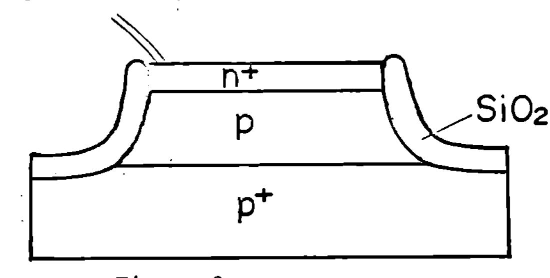


Figure 8 :
Ion implanted APD.

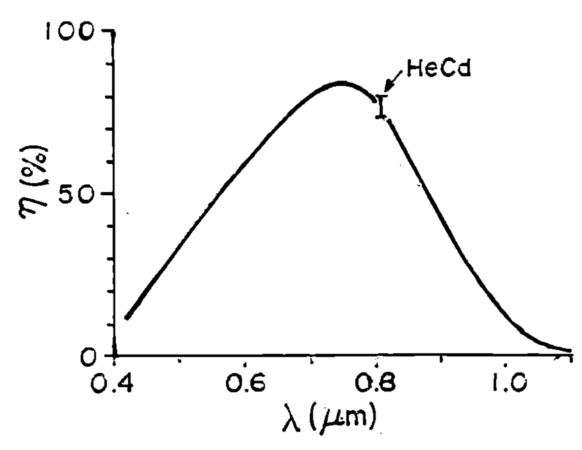


Figure 9 :
Spectral response of an APD

The junction capacity and dark current at 30 volt bias voltage are 1.5 pF and 30 pA, respectively. The noise multiplication factors x are 2.5 and 2.6 at 1.06μ and 0.8μ , respectively.

2) *Stabilization of APD operation*

The characteristics of ADP are so sensitive to temperature variation that a compensation circuit is usually annexed to an optical receiver. As one example, the dc-bias voltage on an APD is controlled by using an electrical circuit involving Zener diodes.

C. Optical fibers as transmission line

1) *Types of fiber and their performances*

Optical fibers are classified into several types with respect to structure and material. While optical fibers are divided into clad type and graded refractive index type as to structure, vitreous silica and glass are two principal mother materials to classify fibers.

Table 3 shows general properties of the two materials for fibers. Optical fibers made of vitreous silica certainly give low loss performance in virtue of its low concentration of impurity. But they possess a high melting temperature property which makes the employment of crucible impossible and leads to the adoption of preform method. Such a non-continuous production process may imply high cost in the future mass production stage.

	MAKING OF LOW LOSS FIBERS	OTHER PROPERTIES
VITREOUS SILICA	EASY	HIGH MELT.TEMP (1800°C) : Possibility of Poor Productivity High Cost.
GLASS	NECESSARY TO PERFORM EXCELLENT MATERIAL PURIFICATION	LOW MELT.TEMP. (1300 °C) Potentiality of Good Productivity Low Cost.

Table III :
Comparison between silica and Glass Fibers.

On the other hand, glass fibers have the potentiality of good productivity, because they can be made at lower temperature permitting one to use such a continuous manufacturing method as double crucible method. But glass fibers obligatorily require an establishment of extremely high grade material purification process to obtain low loss characteristic.

Table 4 describes a rough comparison between various kinds of fibers. Silica fibers made by Corning [11] and Bell Laboratories [12] give 2dB/km loss. A SELFOC® fiber made of borosilicate glass give 14 dB/km loss as indicated in Fig 10, and it will be certain to give much lower loss in future.

As to available band-width, graded refractive index types are superior to multimode clad types. Figure 11 shows experimental data of pulse broadening through a SELFOC® fiber with controlled refractive index profile. Pulse broadening was less than 0.5 nsec/km.[13] Although a single-mode fiber predominates others in available band-width, it renders manufacturing and splicing difficult.

2) *Cabling and splicing*

Glass fibers are susceptible to external forces, particularly to shear force. A plastic coating on glass fibers makes the handling of fibers easy. For field use, a fiber bundle with or without tension members must be protected by some kinds of metal sheaths. In any case, it seems that the cabling technology for fibers is not very much different from that for electrical cables.

Splicing is also very important problem from the viewpoint of practical installation. The average distance between man-holes in cities is on the order of 250 m. This suggests that there will be many connectors between two repeaters. Consequently, even if a slight coupling loss occurs at each connection point, total loss between repeaters becomes serious. Besides, since fiber connection is not laboratory work but field work, the fiber connection work using a laser beam seems difficult to apply to field use. Accordingly, all ingenious and trouble some preparation for fiber connectors must be made in advance in plants and quick and simple works only should be done in man-holes. Furthermore, it is to be noted that some connectors will be employed in an optical transmitter-receiver.

3) *Installation and maintenance*

In addition to optical fiber cable development it must be studied how to install them, whether a special technology is necessary or not in this installation, how to detect a broken fiber point, and so on. This sort of art should be established in the near future.

▪ V. SYSTEMS

A. Types of optical modulation

In general, optical modulation must be made in intensity, because all practical photo-detectors are for intensity detection. But the modulation linearity of DH lasers and the detection linearity of APD are not good. In order to reconcile the two facts, the double modulations such as intensity modulation by a frequency modulated subcarrier signal (FM-AM) and intensity modulation by a PCM or PPM signal (PCM-AM or PPM-AM) are preferably employed in optical communications. The intensity modulation by a pulse signal (PCM, PPM) generally appears better than that by a FM subcarrier signal from the considerations of the DH laser degradation

	SELFOC Fiber (New Type)	Clad Type Glass Fiber: Single mode Fiber	Clad Type Glass Fiber: Multimode Fiber	Single Material Fiber (BTL)	Eccentric Core Optical Fiber (NTT)
Structure	Glass; n; o; ~200μ; Core (30~50μ)	Glass; a; n; o; ~2μ; 50μ; 100μ	Glass; n; o; ~50μ	10	(SiO_2–B_2O_3); SiO_2
Preparation	Ion-Exchange	Double Crucible / Corning Method	CVD	Heat-stretching	Heat-stretching
Characteristics: Bandwidth	Ultra-wideband	Ultra-wideband	a few 10 MHz		
Characteristics: Trans. Loss	~14dB/km at 0.8μ	? ~ 20 ~ 100 dB/km (Corning)	2dB/km at 1.06μ BTL	~5 dB/km at 1.06μ	10 dB/km at 0.8μ
Characteristics: Connection	Easy	Difficult	Easy	Medium	
Characteristics: Mechanical Strength	Good	Good	Good	Good	Good
Characteristics: Degradation	Small	Small	Small	Small	Small
Characteristics: Mass Productivity (Cost)	Good (Continuous)	?	?	Bad	

Table 4:
Comparison between various types of fibers.

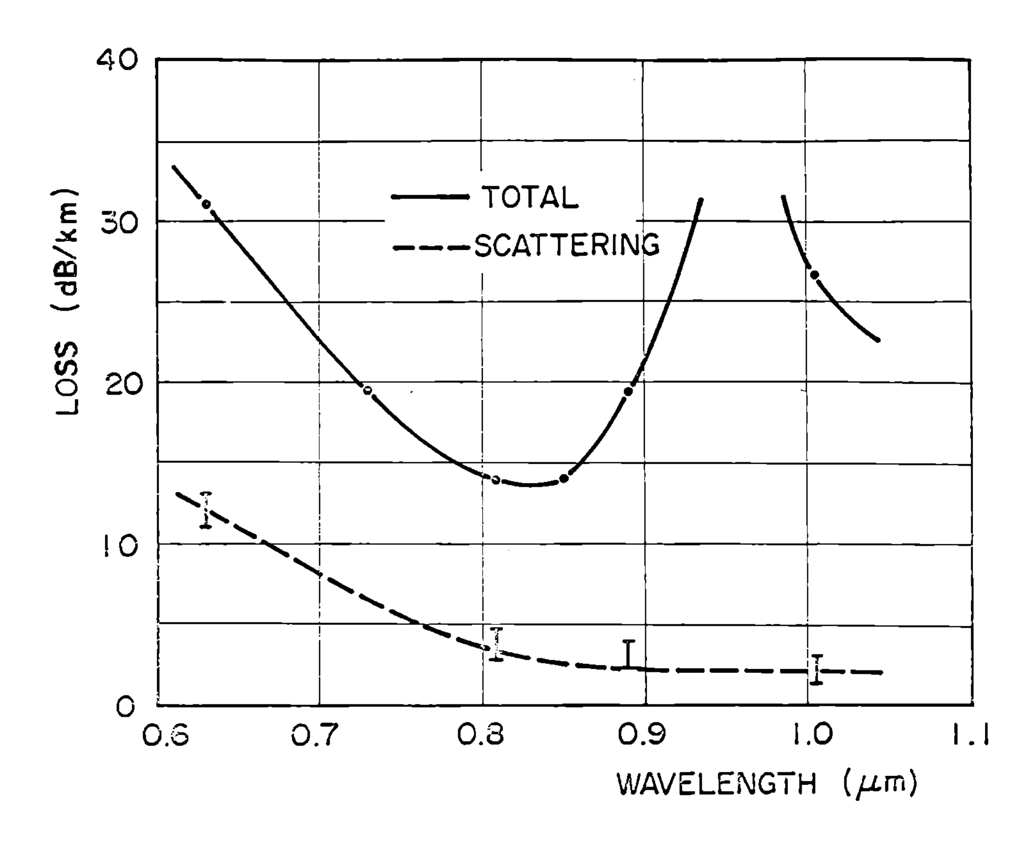


Figure 10 :
Measured loss of a new Selfoc fiber.

due to drive current and allowable span loss.

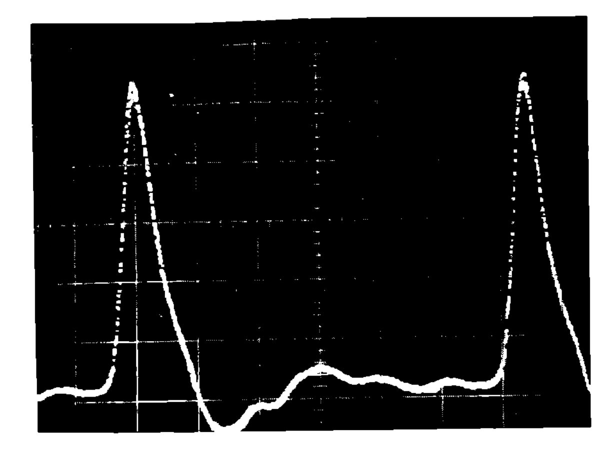
More complicated modulation systems will naturally be investigated in future stage.

B. Types of relay system

Two types of relay systems are generally conceivable with respect to optical fiber communications. One is to employ an electrical amplifier combined with an APD and DH laser. Another is a direct repeater system using optical amplifiers. The former has been naturally investigated more thoroughly compared with the latter. Since an equalization of frequency response against temperature variation is unnecessary in fiber communication systems, a repeater using a regenerative or non-regenerative electrical amplifier becomes very simple in structure.

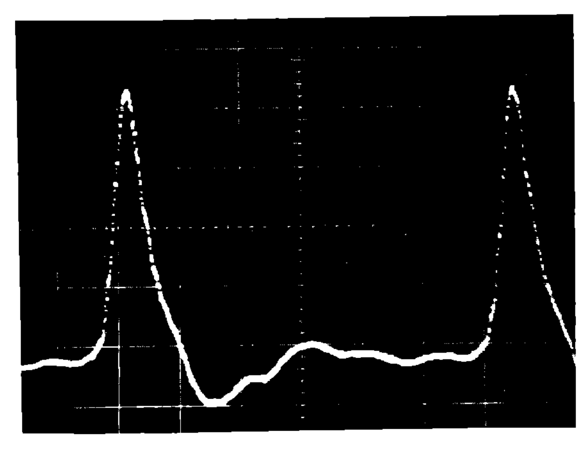
C. Experimental systems

Figure 12 shows an experimental equipment using one kilometer SELFOC fiber. A DH laser and an APD are employed as optical sources and photo-detector. The output of the DH laser is amplitude-modulated by a 20 Mbit/s delta modulated signal of monochromatic TV.



Input Optical Pulse

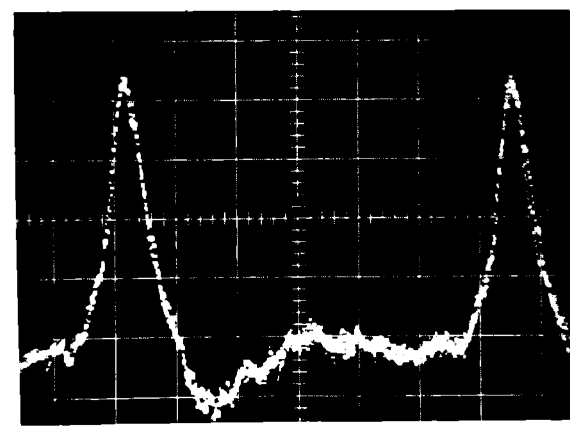
(0.5ns/div)



Output Optical Pulse :

On-Axis Launching

(0.5ns/div)



Output Optical Pulse :

10μm Off-Axis Launching

(0.5ns/div)

Figure 11:
Pulse broadening through a Selfoc fiber with good refractive index profile (fiber length of 320 m and mode-locked Nd:doped YAG laser).

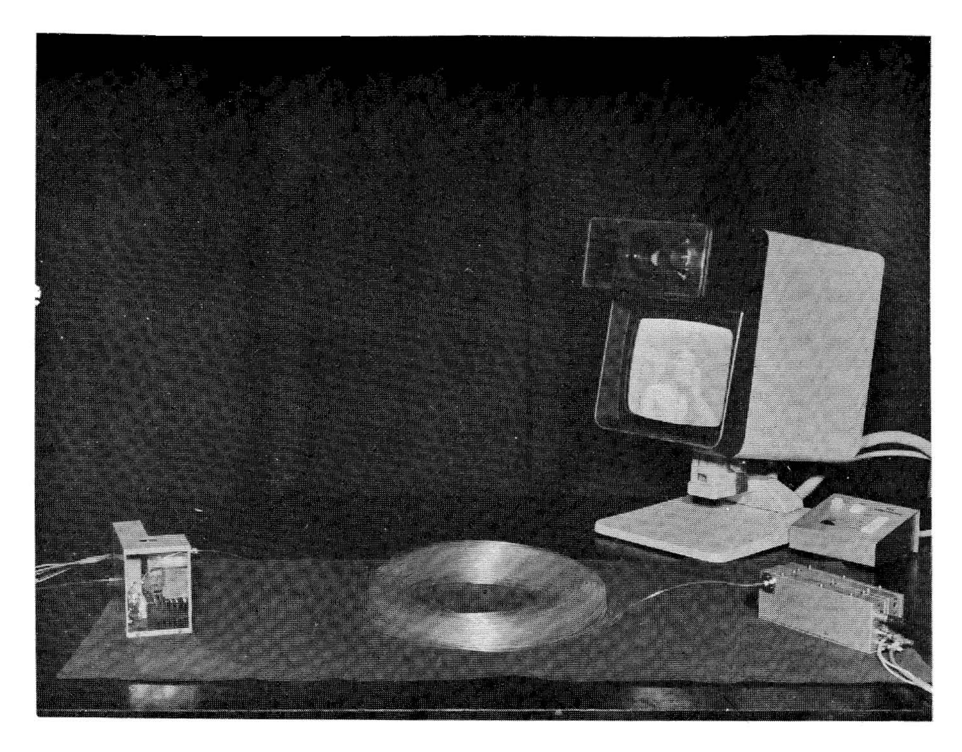


Figure 12 :
An experimental fiber communication set-up of 1 km length.

Figure 13 gives the measured data of bit error rate versus optical receiving power in case of a 100 Mbits/s fiber transmission system. A slight difference between theoretical and experimental curves is observed.

VI. CONCLUSION

The state-of-art concerning optical fiber communication systems has been developed. By making ingenious use of each feature described in Part II optical fiber communication systems will certainly be realized as a wide variety of operating systems such as a broad band long haul system as well as short haul, on-premise, and other systems. It is, however, to be remarked that many essential problems still remain to be resolved : for example :
(1) DH lasers : improvements of life, output power, modulation and coupling characteristics, etc.

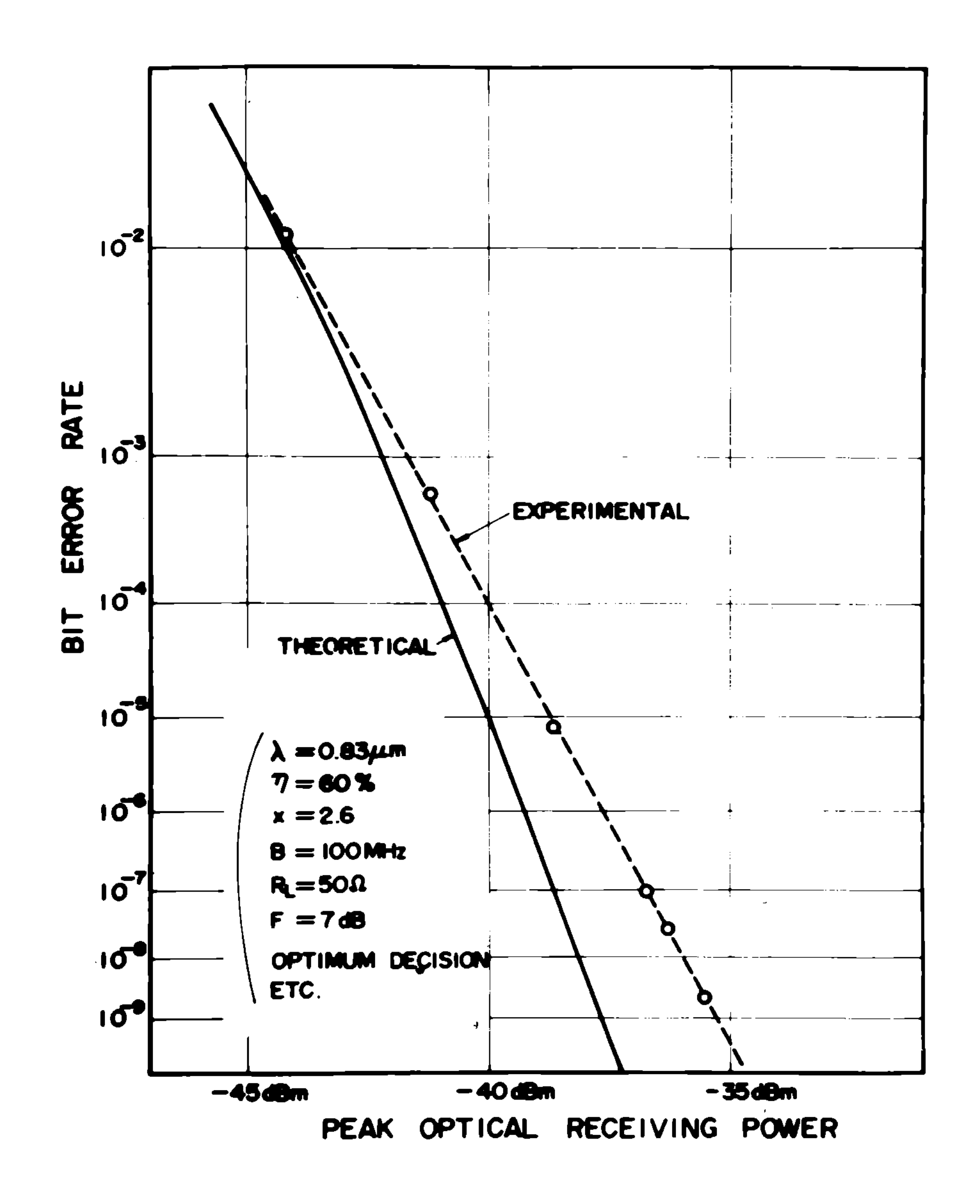


Figure 13 :
Bit error rate vs. optical receiving power (100 Mb/s).

(2) Optical fibers : improvement of characteristics such as loss and pulse broadening , establishment of mass production technology for fiber, establishment of total technologies involving cabling, splicing , installation and maintenance, etc.

REFERENCES

[1] C. Kao et al : Proc. I.E.E.E. 113, pp. 1151-1158 July 1966.

[2] T. Uchida et al : I.E.E.E. J. Quantum Electronics, Vol QE-6 pp. 606-612, Oct. 1970. Press Release Nov. 1968.

[3] B.C. DeLoach : Conference on Laser Engineering and Applications (CLEA), s 12-9, May 30-June 1, 1973.

[4] I. Hayashi : Ibid ., s 12-8

[5] S.E. Miller et al.: Proc. I.E.E.E. , Vol 61, n° 12 p. 1703 Dec. 1973.

[6] R.B. Chesler et al : Appl. Phys. Lett., Vol 23, p. 235 Sept. 1973.

[7] T. Uchida : CLEA s 3-2, May 30 -June 1 , 1973

[8] K. Kobayashi et al : to be presented at 1974 National Convention of the Institute of Electronics and Communication Engineers of Japan (IECEJ), July 1974.

[9] K. Nishida : Japanese Journal of Appl. Phys., Vol 9 n° 5, p. 481 May 1970.

[10] M. Nakajima et al : to be presented at 1974 National Convention of IECEJ , July 1974.

[11] P.C. Schultz : reported at 1973 American Ceramic Soc. Meeting, Cincinnati .

[12] W.G. French et al : Bell Sys. Tech. J., Vol 53, n°5, p.951 May-June 1974.

[13] T. Sumimoto et al : to be presented at National Convention of IECEJ, July 1974.

PROPAGATION DE LA LUMIERE DANS LES FIBRES OPTIQUES

M. TREHEUX, A. COZANNET, R. BOUILLIE

C.N.E.T. 22301 LANNION, France

▪ INTRODUCTION

Il n'est pas question dans un exposé de ce type de résumer toutes les études faites dans ce domaine, à la fois vaste par les différents types de fibres et par les méthodes que l'on peut utiliser pour rendre compte des divers phénomènes. Nous nous attacherons essentiellement à justifier dans une premier partie les études faites vis-à-vis des systèmes qui commencent à être étudiés, et à déterminer les paramètres importants qui interviennent en fait dans une transmission d'informations numériques.

Une seconde partie donnera un aperçu des méthodes que nous avons employées pour étudier la propagation dans différents types de fibres.

Une troisième partie concernera l'aspect expérimental et comparera les résultats acquis aux différentes méthodes de calcul. En conclusion, on examinera l'application de ces techniques à un système de transmission numérique à 8,448 Mb/s.

▪ I. LES FIBRES OPTIQUES UTILISEES EN TRANSMISSION NUMERIQUE

Si nous envisageons un système classique de transmission numérique sur fibres conforme au schéma 1, nous avons entre l'entrée électrique et la sortie électrique un élément optoélectronique que nous pouvons assimiler à un quadripole. Ce quadripole réalise en général une transformation linéaire. Il est certain que pour le milieu de propagation, peu d'expériences ont démontré le contraire (en dehors de l'apparition d'effet Brillouin stimulé à très forte puissance optique). Pour les autres éléments (modulateurs, détecteurs), il est toujours possible de rendre compte de leur non-linéarité en dehors du quadripole optoélectrique e(t), s(t).

Compte tenu de ces remarques nous pouvons supposer qu'entre e(t) et s(t) nous avons la relation de convolution classique.

$$s(t) = e(t) \otimes r(t)$$

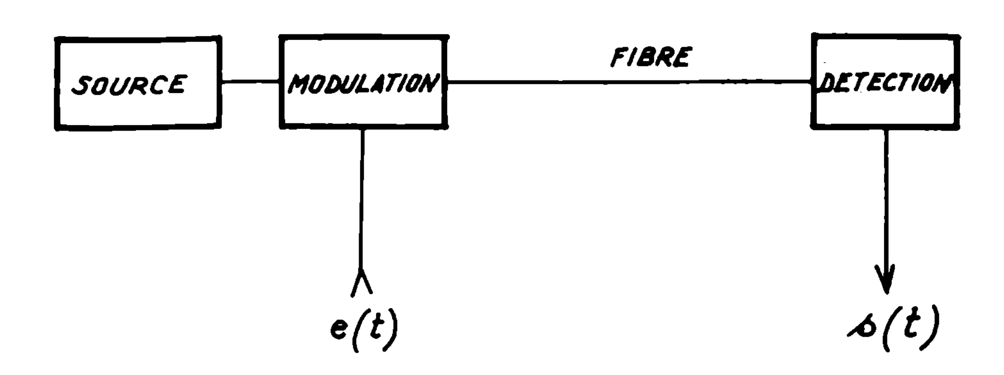


Figure 1 :
Système de transmission sur fibre

où r(t) est la réponse impulsionnelle du système.
Or en considérant à présent ce système comme devant transporter une information numérique on sait, d'après les résultats des transmissions sur câble [1] que le signal à régénérer avant décision numérique donnant le meilleur taux d'erreur est de la forme :

$$f(t) = A \frac{\sin \pi F t}{\pi F t}$$

où F est la fréquence de rythme du train numérique. (Fig. 2).
Partant du signal s(t) de sortie pour réaliser le signal f(t), il faut faire intervenir un filtre dit filtre d'égalisation. La fonction de transfert de ce filtre sera donnée par l'équation :

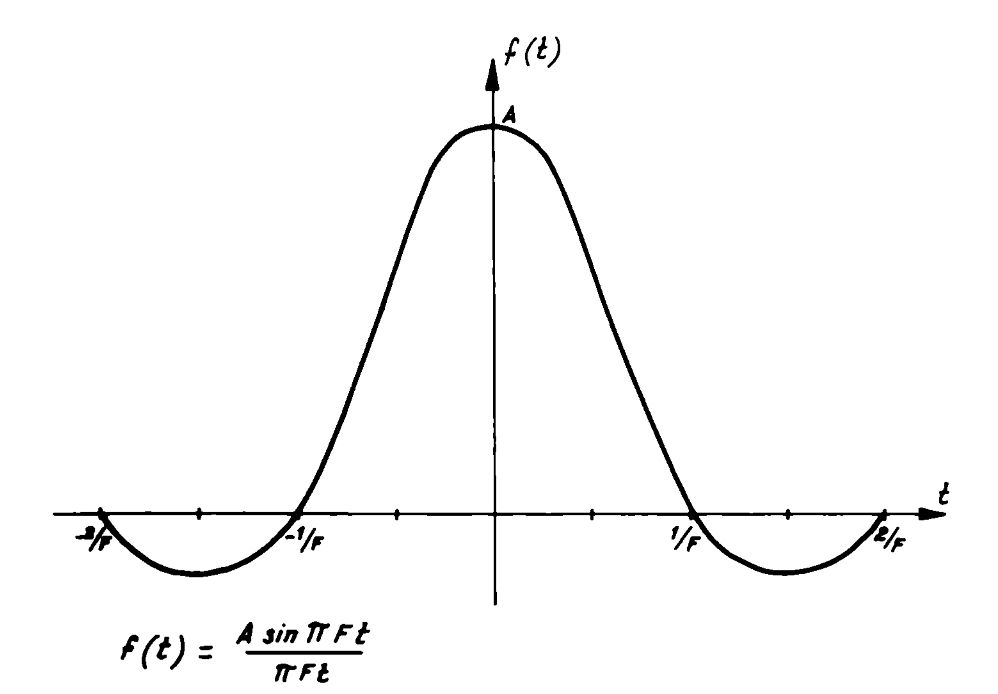
$$F(\Omega) = S(\Omega) \times H(\Omega)$$

où F(Ω) et S(Ω) sont les transformées de Fourier de f(t) et s(t)
Or $S(\Omega) = E(\Omega) \, . \, R(\Omega)$, $F(\Omega) = \pi_{F/2}(\Omega)$

$$\text{Soit : } H(\Omega) = \frac{\pi_{F/2}(\Omega)}{E(\Omega) \, . \, R(\Omega)}$$

On peut par ailleurs décomposer E(Ω) en deux :
♦ l'une correspond au signal d'entrée réel $E_0(\Omega)$, ♦ l'autre à la fonction de transfert du quadripole modulation M(Ω). De même R(Ω) peut être décomposée en deux parties : une fonction électrique liée à la détection et au préamplificateur de détection D(Ω), l'autre étant liée essentiellement à la fibre $R_i(\Omega)$;
- L'équation plus complète devient alors :

$$H(\Omega) = \frac{\pi_{F/r}(\Omega)}{E_0(\Omega) \, M(\Omega) \, R_i(\Omega) \, D(\Omega)}$$



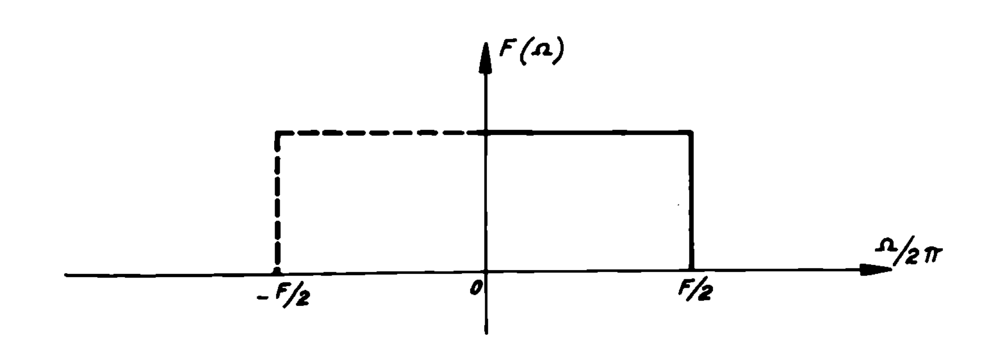


Figure 2 :
Signal à régénérer avant décision en techniques numériques.

Nous voyons donc sur cette formule que la réalisation d'un système numérique nécessite de connaître correctement les différentes fonctions de transfert E (Ω) ; M(Ω) ; D(Ω) ; $R_i(\Omega)$.

Nous verrons dans la troisième partie comment les déterminer expérimentalement. Il serait possible en fait de ne pas se poser de question et de mesurer directement S(Ω) et de réaliser le filtre correspondant. Toutefois étant donné les méthodes de mesure, les limites de détection ne seront pas les mêmes dans le cas du système où l'on pourra obtenir des seuils très inférieurs. Il sera donc nécessaire de faire une extrapolation qui ne peut se réaliser qu'en déterminant chacune des fonctions précédentes.

Par ailleurs nous disposons d'un certain choix sur les différents éléments et il est intéressant de pouvoir faire varier les paramètres. Parmi ceux-ci, les fonctions E(Ω), M(Ω) et D(Ω) sont assez faciles à déterminer expérimentalement. C'est pourquoi nous allons essayer de rechercher les formes de $R_i(\Omega)$ pour les différentes fibres et ses variations en fonction de la longueur de la fibre.

▪ II. ETUDES THEORIQUES DE LA PROPAGATION DANS LES FIBRES

Les différentes études portent en général sur la détermination de la réponse impulsionnelle qui, par transformée de Fourier, permet d'atteindre $R_i(\Omega)$. Dans les premières études réalisées, nous avons considéré le cas de la propagation dans l'approximation de l'optique géométrique, ce qui est logique pour des fibres surdimensionnées par rapport à la longueur d'onde (multimodes, Selfoc, etc.). Ce modèle ne tient pas compte de la diffusion qu'il serait toutefois possible d'introduire dans une théorie plus complète.

A. Calcul général

La base de cette théorie est l'équation générale de propagation d'un rayon lumineux dans un milieu d'indice quelconque. [2]

$$\frac{d}{ds}\left(n \frac{d\vec{r}}{ds}\right) = \vec{\text{grad}}\,(n)$$

où n(x, y, z) est l'indice du milieu .

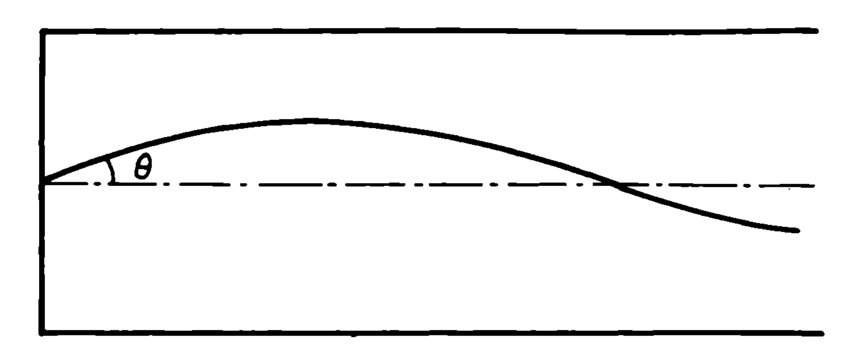
♦ $\vec{r}$ est le rayon vecteur de la trajectoire

♦ s est l'élément d'arc de la trajectoire

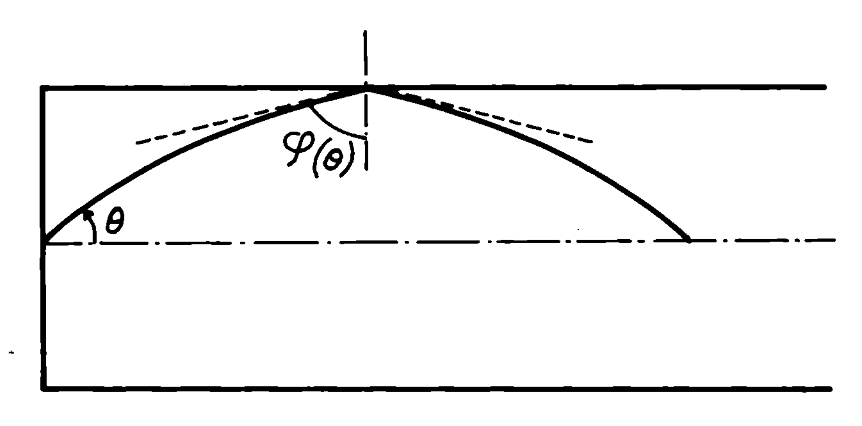
Cette équation est valable quel que soit le milieu (en particulier pour un milieu à gradient d'indice) et pour tout rayon d'entrée (en particulier les rayons obliques). Dans le cas classique des fibres de gradient transverse qui se conserve le long de l'axe Oz, il est normal de transformer l'équation de départ en coordonnées cylindriques.

Un tel calcul n'est pas simple et les possibilités de résultats analytiques sont limitées. Toutefois, nous l'avons utilisé dans deux cas où il est possible : celui des fibres multimodes et des fibres à gradient d'indice de la forme :

$$n^2 = n_0^2\,(1 - \alpha^2\rho^2)$$



a) Propagation sans discontinuité



b) Propagation avec discontinuité (définition de φ(θ))

Figure 3 :
Propagation de la lumière dans une fibre à gradient d'indice quelconque.

Une publication en cours donnera les résultats correspondants. On peut dire que dans ces deux cas précis, les réponses impulsionnelles obtenues sont soit identiques (fibres multimodes), soit très voisines (fibres à gradient quadratique) de celles que l'on obtiendrait dans le cas de rayons méridiens. Cependant une nouvelle notion d'angle d'acceptance apparait qui permet d'introduire de la

lumière sous des angles plus grands que ceux qui sont définis habituellement par l'ouverture numérique.

B. Propagation dans les fibres pour une injection à symétrie de révolution autour de l'axe.

Considérons une fibre quelconque et un rayon méridien d'entrée faisant avec l'axe l'angle θ. Dans la mesure où nous négligeons la diffusion, entre la durée du trajet sur une longueur L et θ, existe une relation certaine $\tau(\theta)$. En supposant que la source émette un cône élémentaire autour de l'angle θ le faisceau arrivant à l'instant t au bout de la longueur L a une énergie donnée par :

$$dE(\theta,t) = e^{-\alpha\sigma(\theta)}\ a(\phi(\theta))^{N(\theta)}\ f(t - \tau(\theta))\ H(\theta)\ \sin\theta\ d\theta$$

où $\alpha\sigma(\theta)$ représente les pertes le long du trajet géométrique, $a(\phi(\theta))^{N(\theta)}$ les réflexions éventuelles sur des surfaces de discontinuité d'indice, f(t) la répartition temporelle de l'énergie d'entrée, $H(\theta)$ la brillance de la source.

On obtient la forme de l'impulsion de sortie par :

$$f(t) = \int_0^{\theta_0} e^{-\alpha\sigma(\theta)}\ a(\phi(\theta))^{N(\theta)}\ f(t -\tau(\theta))\ H(\theta)\ \sin\theta\ d\theta$$

où θ_0 est l'angle d'acceptance de la fibre.

On remarque que l'angle $\phi(\theta)$ sous lequel est attaquée la surface de discontinuité se conserve tout au long de la fibre (Loi de Descartes). Etant donné la relation certaine $\tau(\theta)$ on peut inverser cette fonction ce qui donne :

$$f(t) = \int_{\tau_0}^{\tau(\theta_0)} e^{-\alpha\sigma'(\tau)}\ a(\phi'(\tau))^{N'(\tau)} H'(\tau)\ \psi(\tau)\ f(t-\tau)\ d\tau$$

en appelant ( )' les fonctions correspondantes après changement de variable. Cette expression s'écrit par ailleurs :

$$f(t) = \int_{-\infty}^{+\infty} \pi_{\tau_0,\tau(\theta_0)}(\tau)\ e^{-\alpha\sigma'(\tau)}\ a(\phi'\ (\tau))^{N'(\tau)} H'(\tau)\ \psi(\tau)\ f(t-\tau)\ d\tau \quad \text{où :}$$

$\pi_{\tau_0\,\tau(\theta_0)}(\tau)$ représente la fonction porte entre τ_0 et $\tau(\theta_0)$. On obtient ainsi d'une manière simple la réponse impulsionnelle cherchée $r_i(t)$ donnée par :

$$r_i(t) = \pi_{\tau_0,\tau(\theta_0)}(t)\ e^{-\alpha\sigma'(t)}\ a(\phi'(t))^{N'(t)}\ H'(t)\ \psi(t)$$

et $R_i(\Omega)$ par sa transformée de Fourier.

Ce résultat est valable quel que soit le type de fibres. Pour le calcul appliqué à une loi précise de l'indice il faut :

- calculer la trajectoire du rayon méridien,
- calculer la longueur géométrique $\sigma(\theta)$ et optique $c\tau(\theta)$ pour une longueur L,
- inverser la fonction $\tau(\theta)$,
- en déduire les fonctions σ', ϕ', N', H' et ψ

1. *Fibres multimodes.*

Le calcul a été fait directement dans ce cas. On obtient les résultats suivants : [3,4]

a-1) Influence du diagramme d'émission $H(\theta)$:

Aussi bien sur $r_i(t)$ que sur $R_i(\Omega)$, ce paramètre influe directement. Dans le calcul précis on a pris une injection de la forme $(\cos\theta)^m$ qui correspond à des sources diverses :

laser classique avec optique	$m = 10^3$ à 10^4
diode laser	$m = 500$ à 5000
diode électroluminescente	$m = 10$

a-2) Influence de l'atténuation :

On montre que si son influence est prépondérante sur le niveau de sortie, elle est quasiment nulle sur la forme de fonction de transfert

a-3) Influence des pertes à l'interface :

les pertes à l'interface apparaissent comme un filtrage de modes et en particulier des modes élevés. On peut les introduire en considérant une loi de réflexion de la forme $a(\theta)$. Suivant que :

$1-a(\theta) = a_0$ ou $a_1\,\theta$
ou $a_2\,\theta^2$
ou $a_3\,\theta^3$

on a les lois représentées sur la figure 5 pour l'élargissement à mi-hauteur Δt. Ces résultats sont inversés pour la bande passante $\Delta\Omega$.

Suivant les théories faites pour essayer de déterminer $a(\theta)$, on peut avoir une loi en θ^2 ou θ^3 , l'expérience séparant difficilement les deux cas.

2. *Fibres à gradient d'indice*

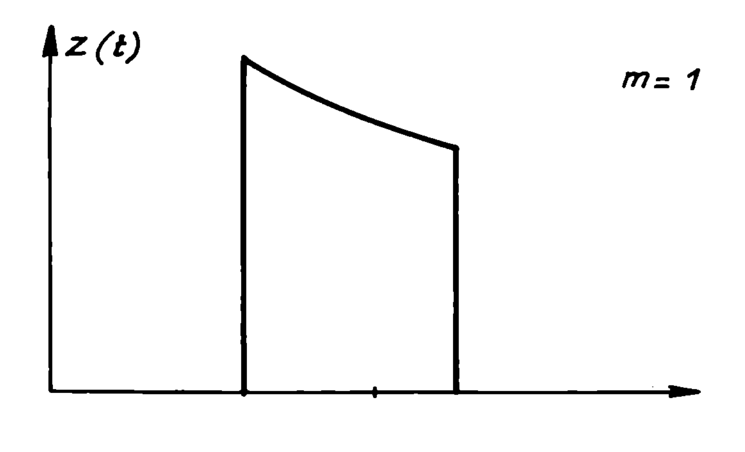
Dans ce cas, il est tout d'abord nécessaire de déterminer les lois $\rho(z)$, $\tau(\theta)$. Pour cela on doit partir de l'équation générale qui se simplifie (plan méridien).

$$\frac{d^2\rho}{dz^2} = \left(1 + \left(\frac{d\rho}{dz}\right)^2\right)\frac{1}{n}\,\frac{dn}{d\rho}$$

soit :

$$\begin{cases} 1 + \left(\dfrac{d\rho}{dz}\right)^2 = n^2 k^2 \\[2ex] \tau = \dfrac{1}{c}\displaystyle\int_0^L k\, n^2\, dz \end{cases}$$

Il est évident que ces équations ne s'intègrent pas pour des profils d'indice quelconques. Toutefois elles s'intègrent pour les profils suivants :





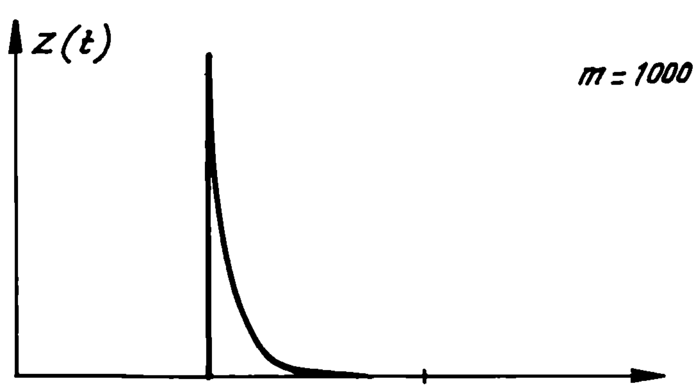
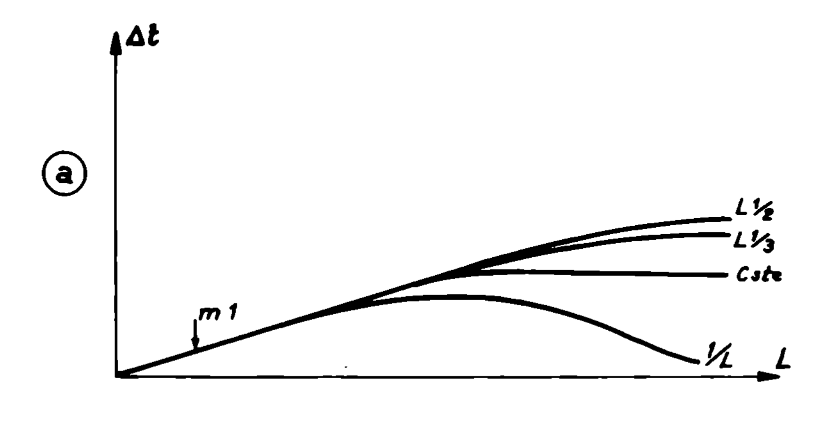


Figure 4 :
Influence de l'injection sur l'impulsion de sortie

a) m =I
b) m = I00
c) m = I000



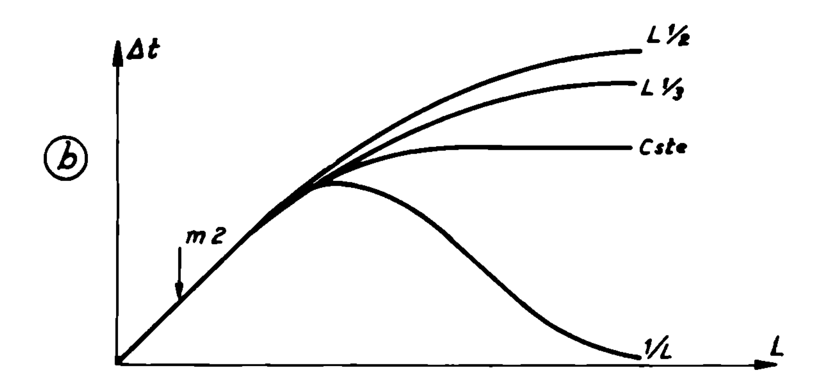


Figure 5 :
Influence de la loi 1-$a(\theta)$ sur l'élargissement à mi-hauteur Δt

a) m_1 *grand*
b) m_2 *petit*

$$n = n_0$$

$$n = n_0 \ (1 - \alpha x)$$

$$n = \frac{n_0}{\mathrm{ch}\,\alpha x}$$

$$n = n_0 (1 - \alpha^2 x^2)^{1/2}$$

$$n = n_0 \ (1 + a \ e^{-bx})$$

Partant des lois ci-dessus, il est possible d'intégrer un profil quelconque en approximant celui-ci par une suite de profils intégrables. C'est ce qui a été dans la référence [5] .

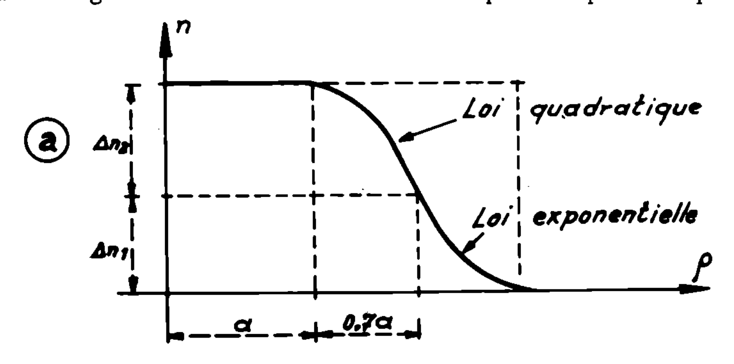
Pour ces profils, le temps de propagation est donné par :

$$\tau(\theta) = L \frac{\sum \Delta\tau_j}{\sum \Delta z_j}$$

suivant la partition réalisée.

L'inversion est délicate à faire directement, cependant par traitement numérique il est possible de faire à la fois l'inversion et le calcul de $r_i(t)$.

Le calcul fait, par exemple, sur un profil d'indice conforme à la Figure 6 où la frontière est à profil quadratique montre une



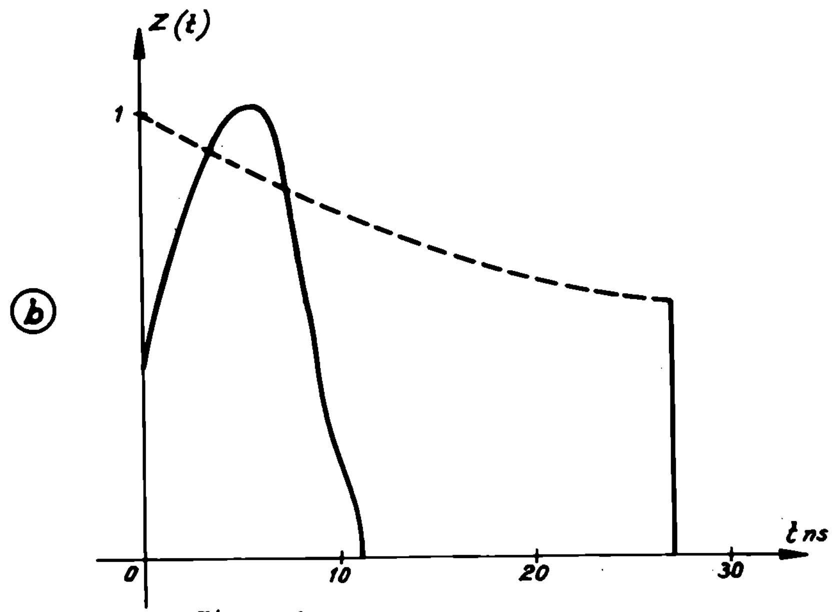


Figure 6 :
Applications du calcul par approximations dans un cas particulier.
a) profil d'indice choisi } ---- *cas multimode*
b) réponse impulsionnelle } ____ *avec gradient*

réduction de plus de 50% de l'impulsion élargie par rapport au profil rectangulaire.

Ce processus a l'avantage de s'appliquer à un profil réel puisqu'il est possible de le mesurer par différentes méthodes (microscope interférentiel, dosage de l'élément créant le profil, etc....)

▪ III. ETUDE EXPERIMENTALE DE LA PROPAGATION DANS LES FIBRES

Le but de cette étude est de déterminer avec une certaine précision la fonction de transfert $R_i(\Omega)$ dans des conditions d'expériences variables. Nous nous sommes tout d'abord attachés à déterminer l'élargissement à mi-hauteur d'une impulsion fine pour les différentes fibres. Ceci a conduit aux résultats suivants :

1. *Influence de l'entrée.*

Cette influence a été démontrée précisément sur différents types de fibres , exemples :

*fibre à coeur liquide	:	variable de 1 ns/km à 200 ns/km
*fibre CORNING	:	" de 1 ns/km à 10 ns/km
*fibre FORT	:	" de 1 ns/km à 10 ns/km

2. *Influence de la position à l'entrée*

Dans le cas des fibres à gradient d'indice, ce paramètre est important et il est par exemple possible de passer de 1,5 ns/km à 5 ns/km dans une fibre SELFOC.

3. *Influence du gradient d'indice.*

Cette influence est certaine puisque, pour une même injection, on peut passer de 10 ns/km (fibre CORNING) à 2 ns/km (fibre SELFOC). Ceci explique les efforts faits par différents laboratoires pour améliorer la réponse des fibres par l'introduction de gradient d'indices.

4. *Influence de la longueur.*

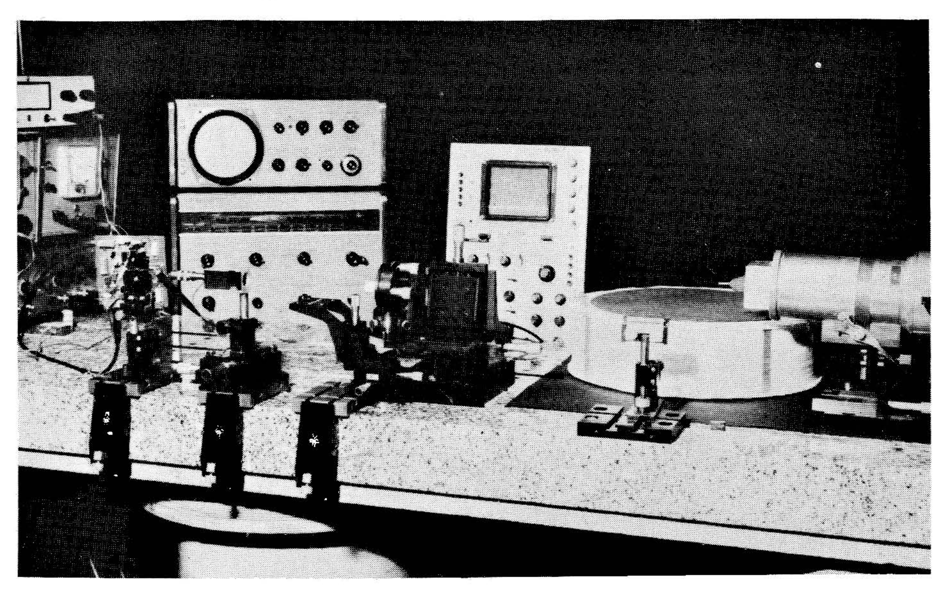
A l'heure actuelle, nous n'avons pas de cas précis où la loi $\Delta t(L)$ ne soit pas linéaire. Exemple : 1, 2, 3 km de CORNING →5, 10, 15 ns/km. L'expérience utilise une source d'impulsions fines (100 ps) formée d'une diode laser, d'un laser à Argon ou YAG à modes bloqués ou éventuellement d'un diode électroluminescente (qq ns).

Un système optique variable permet de réaliser des conditions d'injection diverses. Un système d'injection stable rend les mesures répétitives.

Le système de détection est dormé d'une photodiode rapide (70 ps) et d'un oscilloscope à échantillonnage.

Pour des études sur des longueurs importantes, la diode laser est beaucoup plus souple car tous les paramètres d'émission sont variables, toutefois les paramètres optiques sont moins bien définis.

Une nouvelle méthode de mesure permet de déterminer directement la fonction $R_i(\Omega)$. Partant d'un émetteur d'impulsions quel-



Photographie n° 1:
Mesure de la fonction de transfert $R_i(\Omega)$ sur I kilomètre.
Photographie CNET No 501894.

conques (il suffit que ces impulsions soient plus courtes que l'impulsion obtenue au bout de la longueur choisie) caractérisées par leur spectre $E_0(\Omega)$, on commence par les détecter après une très faible longueur de fibre, à l'aide d'un détecteur (de fonction de transfert $D(\Omega)$). On mesure ainsi sur un analyseur de spectre la fonction $|E_0(\Omega)\ D(\Omega)|$. On réalise ensuite le système complet avec une longueur L de la fibre et l'on mesure alors :

$$|E_0\ (\Omega)\ R_i(\Omega)\ \ D(\Omega)\ |$$

En utilisant des échelles logarithmiques sur l'analyseur de spectre, nous obtenons d'une part :

$$s_0(\Omega) = \log_{10}|E_0(\Omega)\ D(\Omega)|$$

et

$$s_1(\Omega) = \log_{10}|E_0(\Omega)\ R_i(\Omega)\ D(\Omega)\ |$$

en enregistrant les deux fonctions, de manière à avoir $s_1(0) = s_2(0)$ (correction de $e^{-\alpha T_1}$ puis en faisant leur différence nous obtenons directement la fonction de transfert $\log_{10}|R_i(\Omega)|$. Les résultats obtenus sont conformes à la théorie précédente. On a en particulier des formes telles que celles qui sont représentées sur la figure 7

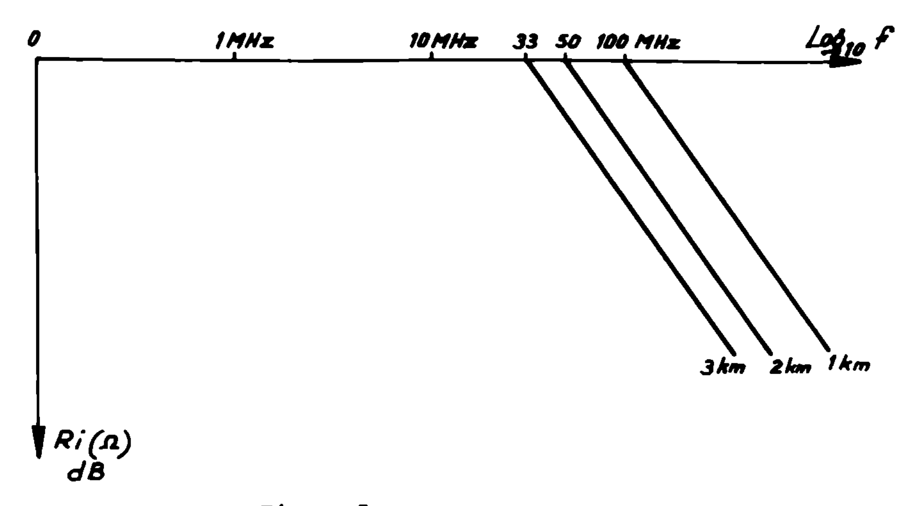


Figure 7 :
Fonctions de transfert d'une fibre multimode pour des longueurs de I,2,3 kilomètres (Plan de Borde).

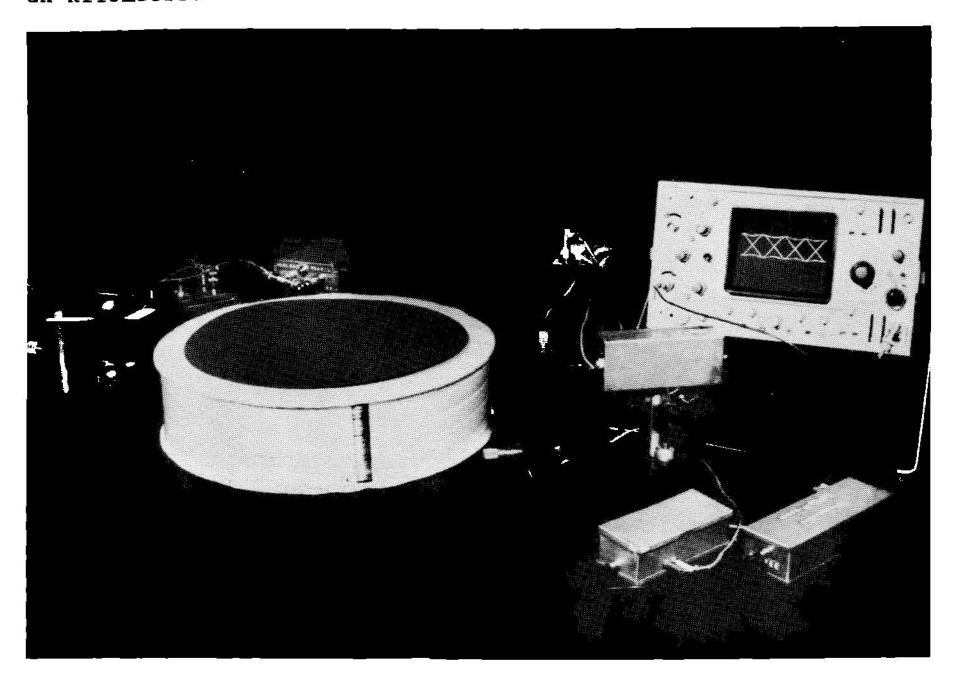
IV. CONCLUSIONS

Nous avons montré que la connaissance de la fonction $R_i(\Omega)$ était fondamentale pour la réalisation d'un système. Nous donnerons comme exemple le cas des fibres CORNING utilisées pour un système à 8,448 Mb/s. La fonction de transfert obtenue après un kilomètre a la forme représentée Figure 7. Par ailleurs on a montré expérimentalement que les fonctions de transfert pour 1,2,3 km se déduisaient par une translation dans le plan de Bode, le produit $\omega_c L$ restant constant. Ceci veut dire qu'avec une fréquence de coupure f_c = 30 MHz pour 1 km , on obtiendrait une f_c = 5 MHz pour 6 km. Comme par ailleurs la demi fréquence de rythme est égale à 5 MHz, ceci veut dire que l'on n'a pas a tenir compte de la fonction de transfert de la fibre jusqu'à 6km pour réaliser l'égaliseur.

Ce cas correspond à une diode laser placée contre la fibre et remplissant largement l'angle de celle-ci. C'est donc également celui d'une diode électroluminescente et il est possible de faire le calcul de l'égaliseur correspondant sur 6 km sans prendre en compte la fibre. C'est ce qui a été fait par exemple dans un article récent [6] .

Par ailleurs, à l'aide de la même diode laser et d'une optique, on a montré que sur 1 km on pourrait avoir une fréquence f_c de l'ordre de 100 MHz. On pourrait donc dans ce cas, si le bilan énergétique le permettait, aller jusqu'à 20 kilomètres à 8,448 Mb/s. On peut évidemment penser utiliser les fibres correspondantes à des rapidités plus élevées (2 km à 140 Mb/s par ex.) Il est certain que dans la mesure où $R_i(\Omega)$ n'intervient pas dans le filtre $H(\Omega)$, celui-ci est beaucoup plus facile à réaliser. Il sert en fait à redresser uniquement le circuit RC de réception correspondant à $D(\Omega)$ que l'on a intérêt à rendre le plus grand

possible pour des raisons liées au rapport Signal/Bruit. Dans les systèmes que nous mettons au point actuellement, c'est ce qui est réalisé. La photographie 2 représente un tel ensemble sur un kilomètre.



Photographie 2 :
Système de transmission à 8,448 Mb/s.
Photographie CNET No 501886.

Nous avons traité ici le problème de la transmission sur fibres en partant du point de vue de l'utilisateur transmettant une information numérique. Il est certain que les théories faites ne sont pas complètes et en particulier il serait nécessaire de travailler dans deux directions. L'une consiste à introduire les phénomènes de diffusion en remplaçant la fonction $\tau(\theta)$ par une relation de probabilité, se rapprochant ainsi des études de transmissions atmosphériques en échos diffus, l'autre nécessite la connaissance expérimentale et théorique des phénomènes de diffusion et permettrait de déterminer cette relation de probabilité. Il reste donc de nombreuses études à faire pour atteindre une connaissance suffisante du milieu de propagation, celles-ci devenant de plus en plus complexes au fur et à mesure de leur développement. Nous tenons ici à remercier l'ensemble du groupe Télécommunications optiques qui a participé à ces travaux et le Docteur Steiner du F.T.Z.*) pour sa collaboration dans le domaine théorique.

*) FERNMELDE TECHNISCHEN ZENTRALAMT FORSCHUNGSINSTITUT 61 DARMSTADT Deutsche Bundesrepublik.

REFERENCES

[1] J. Boulvin, L'onde Electrique 1974 Vol 54 n°2 p.73-80.

[2] Born and Wolf, Principles of Optics, Pergamon Press 1964 p. 122.

[3] M. Tréheux, R. Bouillie, K. Steiner, Annales des Télécommunications Juillet 1974, 29,195,1974.

[4] K.M. Steiner, M. Tréheux and R. Bouillie, Applied Optics Vol 12 n°11 2732-2735, 1973.

[5] A. Cozannet, M. Tréheux , R. Bouillie , Annales des Télécommunications, Juillet 1974,29,219,1974.

[6] C. Boisrobert, A. Leboutet, Annales des Télécommunications Juillet 1974,29,227,1974.

MATERIALS FOR LOW LOSS OPTICAL FIBERS

J. Ernest

Laboratoire de Marcoussis

Departement Recherche Physique de Base

91460 MARCOUSSIS France

INTRODUCTION

The original suggestion of Kao and coworkers[1], in the late 60's, that optical fibers *could* be made with an attenuation low enough to make them suitable for optical communications, is mostly responsible for triggering work in this field, and has resulted in the tremendously increasing effort which we are witnessing today.

Over the last few years, progress has been very fast indeed, glass materials of different types (high silica compositions, silicates and borosilicates) have been prepared with attenuations less than 20 dB/km, a value which was originally set as a target goal for making "communication-grade" optical waveguides. Record performances in the range of 1 to 4 dB/km have been recently announced[2][3]for high silica fibers, and losses for silicate and borosilicate glasses lie in the range 15 to 20 dB/km, with further progress still ahead. An additional remarkable feature of these fibers is that their lowest loss figures lie in the spectral range where suitable optical sources for telecommunications are found, namely 800 to 900 nm (gallium arsenide LED's and laser diodes) and 1.06 μm (Nd doped crystal lasers). The general status of optical communications is reviewed elsewhere in this book.[4]

In this paper, we shall first review the general requirements for an optical fiber material, then proceed to examine the physical origin of optical attenuation in glass materials,while trying to assess theoretical limitations and lastly we shall discuss and compare results obtained with present day best materials and give some indications on further contemplated progress.

I. GENERAL REQUIREMENTS FOR OPTICAL FIBER MATERIALS.

Besides exhibiting low losses in the wavelength range of interest, materials for optical fibers must meet specific requirements that eventually limit the choice in compositions. It is well known that most of the required material properties are fulfilled by glasses of different types. It is true that liquid-core fibers [5] have been fabricated in several laboratories, that do exhibit attenuations as low as 8 dB/km. However, the liquids used, mostly organic compounds, have poor performance in terms of stability versus temperature and durability, and would prove to be very awkward in the field with respect to cabling and splicing problems. This type of fiber is at present not envisaged as a possible candidate for large scale industrial applications, although they have proved useful for testing propagation models in laboratory experiments [6] .

The main properties required for optical waveguide materials are the following:

a. Their refractive index must be modifiable in a range that covers the required core/cladding index differences (for "step-index" fibers) and for radial index profile (for "graded-index" fibers), through material changes that do not deeply affect mechanical and thermal properties. This is because big differences between core and cladding material properties would result in overly large strains in the fiber, while it is being drawn from the "preform". Required maximum index differences (between center and periphery of the fiber) are in the range $\Delta n_{max} \sim$ a few 10^{-3} to a few 10^{-2}, depending upon systems considerations, as was shown in another talk in this conference [7] . Typical desired values are 5×10^{-3} to 2×10^{-2} , for multimode "step" or "graded-index" (SELFOC-type) fibers.

Such index differences may be achieved through several techniques, depending upon the base material and the glass and fiber fabrication processes. Most used techniques are:

- doping with high optically polarizable oxides, like TiO_2, GeO_2 ("index raising") or B_2O_3 ("index lowering") as in the case of glasses with a high silica content (> 90%) [8] .
- changes in the proportions of the components, as in the case of more conventional three (or more) components glasses, like sodalime silicates (SiO_2 - Na_2O - CaO) [9] .
- alkali-ion exchanges, obtained during the fiber drawing process itself, as in the case of borosilicate glasses (exchange between Na^+ and Tl^+ ions), which specifically result in desired index graded profile, like for SELFOC type fibers[10] .

Of course, great care has to be taken, while proceeding to those material modifications, in order to avoid any increase in absorptive impurity content.

b. Characteristic melting, fining, as well as fiber-drawing temperatures should preferably be low, for improving ease of fabrication. Also, at least in the case of glasses of conventional composition (silicates, borosilicates, etc...), low processing temperatures are favorable in terms of reducing contamination of the materials through diffusion of impurities from furnaces, crucibles, and

so on.
c. Materials should be as much as that is compatible with other desired characteristics, "true glasses", i.e., possess a rather wide softening temperature range, so as to avoid excessive straining and/or devitrification (crystallization) while cooling the melt and/or drawing the fibers.
d. Fibers should possess good mechanical properties, i.e., good flexibility and high tensile strength, as well as good resistance to the appearance of microcracks.

All these desired properties are interrelated in some way, which means that it is difficult to reach satisfactory trade-offs. For instance, fibers which exhibit, to date, the lowest reported loss figures, i.e., made of "high silica" glasses, are not necessarely the best with respect to mechanical properties and resistance to microcracks.
e. Finally, the maximum optical loss requirements may appreciably differ, depending upon the types of systems that one has in mind. Generally speaking, however, the most important systems now envisaged, i.e. links of several kilometers length between transmitting and receiving ends, and/or between repeaters, call for optical cables with losses well below 20 dB/km, possibly less than 10 dB/km, at the wavelength ranges of interest (800-900 nm and 1.06 m), as clearly shown by system design considerations, of which T. Uchida gives significant examples at this conference [4]. However, some specific applications, i.e. short range links for data transmission on board ships and airplanes, or within buildings for instance, may call for less stringent requirements: from 20 dB/km up to 50 dB/km may be acceptable values in some cases. In other words, there still seems to be room for types of fibers other than those reported for having the lowest loss figures ("high silica"), provided that the use of different (lossier) materials brings about additional advantages, like high flexibility, a wider range of refractive index differences, and the like. On top of that, one has to bear in mind that systems designers specify maximum loss figures for an optical *cable*. As a consequence, safety factors should be duly taken, when specifying the corresponding maximum loss for the optical *fiber*, as it is now well known that handling of fibers during the cabling process may lead to a possibly appreciable increase in (scattering) loss.

II. THE ATTENUATION PROBLEM

A discussion of the attenuation problem for today high quality fiber materials imply a close inspection of the separate physical causes of the optical loss. In some instances, it is found that a particular contribution to loss, like for instance intrinsic scattering, is experimentally very close to its theoretical limit, in as much as the latter can be assessed with some degree of confidence. In turn, experimental investigations on ultra low loss fibers bring about new results that may help refining existing theories, and shed new light on detailed spectroscopic features of materials, as was shown, for instance, by Keck et al. [4].

Grossly speaking the causes of optical loss fall into two

broad categories:
- absorption, i.e., photons are converted into heat through various impurities, material vacancies, intrinsic absorption bands.
- scattering, i.e., photons will eventually escape guided modes after hitting regions with index inhomogeneities, material defects, etc...

Both processes lead, in principle, to an exponential decay of the guided light intensity, as it travels along the fiber. However, detailed features of scattering loss versus distance along the fiber are usually more complicated, as the fiber exhibits either long-range (curvatures, twists) or short-range (bubbles, cracks, index variations, etc...) deviations that are far from having a regular distribution along the length. In this discussion, we will limit ourselves to the contribution to scattering of the fiber-core material proper. Obviously, the fiber drawing process always contributes to some increase in scattering loss. However, newly developped drawing techniques, either of the "preform" type or the "double crucible" type, tend to largely minimize this contribution. In the best fibers, scattering loss appears then to be mainly due to material scattering. Good reviews of the optical attenuation problem in glass materials are found in[17] and[18].

A. Absorption losses

Until recently, most optical glasses, even among the best, exhibited losses in the range of a few 10^2 to a few 10^3 dB/km, that were mainly attributed to absorption through impurity ions. Technological improvements to reach present day figures of about 10 dB/km, have been heavily aimed at reducing impurity contents. Other, more subtle contributions to absorptive loss are now appearing of importance, i.e. "intrinsic absorption", which might be considered as an ultimate limitation for nearly perfect materials. We will examine these different causes for absorption in turn.

1) Intrinsic absorption

Intrinsic absorption in a nominally transparent glass (in the visible range) is related to the fundamental absorption edge, usually located in the UV, which is similar to the interband gap of semiconductor materials. This interband absorption may be assigned to oxygen atoms, which are major constituents of most transparent (oxide) glasses. According to the Dow and Redfield model[11], exciton levels, located close to the fundamental band edge, are broadened by variations of local electric fields, on a microscopic scale, which is characteristic of the random structure of liquid and glassy states. This results in an absorption tail that may extend far into the visible region, with an exponential dependence on photon energy, of the type:

$$\alpha_{a,i} \propto \exp((E - E_g)/ \Delta E) \qquad (1)$$

Here E is the photon energy, E_g the band edge energy and ΔE is a parameter characteristic of the material. Qualitatively, ΔE may be related to the mean amplitude of the local electric field (and to microscopic density) fluctuations in the material. These fluctuations result from thermal fluctuations existing in the molten or softened material, that are eventually frozen into it, while

it is being cooled to reach the metastable glassy state. Usually, fluctuations are related to a fictive temperature T_f , which is closely related to the softening temperature . Empirical fits of experimental absorption data, for different types of glasses, to the exponential dependence [1] tend to show that ΔE is directly proportional to kT_f.

In other words, one can sum up the way intrinsic absorption will show up, when one compares two different materials, according to this crude model, in the following way : the intrinsic absorption α_{ai} , at a given wavelength (in the visible or near infrared part of the spectrum) will tend to be :

- larger for materials with higher T_f'S (or higher softening points).
- smaller for materials with larger fundamental band-edge energies Eg.

The former indicated trend tends thus to favor materials with comparatively low characteristic melting viz-softening temperatures, (e.g. three or more component-glasses), while the latter indicated trend is in favor of "single-oxide" materials (like high silica glasses), which usually have higher band-edge energies.

Table I sums up a comparison, in this respect, of two types of glasses widely used for optical waveguides, i.e. : pure (or "high") silica glass and soda-lime silicate glass (SiO_2 -Na_2O- CaO).

Type of glass	Eg(eV)*)	ΔE(eV)	α_{ai} (λ≃900 nm)
SiO_2	≃ 5	0.5	≃ 1 dB/km
SiO_2-Na_2O-CaO	≃ 3	0.3	≃ 6 to 7 dB/km

Table I *)

*) After [11]. Eg values are very approximate values for best fits.

From Table I, we could draw preliminary rough conclusions about ultimate (i.e. intrinsic) absorption loss evaluation for two different types of glass that are of prominence today for working optical waveguides : it would seem best to look for materials with highest energy fundamental absorption band edge, as differences in E_g seem to be a more sensitive factor than differences in ΔE. In this respect, pure (or "high") silica glasses are apparently in the best position, as SiO_2 has its U.V. absorption band at shorter wave-

lengths than most other glasses. Indeed, it is true that lowest-loss high silica waveguides published to date [2] exhibit residual absorption losses, at around 800 nm, that are of the order 1dB/km (or may be even less!). However, estimation of residual (intrinsic) absorption loss for other glasses, like soda-lime silicates, is presently obtained by taking and extrapolating data on materials where *impurity absorption losses* are still comparatively high. So that estimates for intrinsic loss, in those cases, are highly doubtful. Only when those more conventional glasses have impurity-related absorption losses as low as those of high-silica materials, will we be able to more safely reassess true intrinsic absorption for them.

2) *Impurity related absorption losses*

As mentioned earlier, impurities in otherwise nominally transparent glasses have been, up to recent times, the predominant cause for optical loss . Impurities are probably the main offenders, at least in the visible range, whenever one observes optical losses even as low as 20 to 50 dB/km.

It has since long been recognized that most important impurities are either transition metal ions (specially of the iron series), which are responsible for most of the varied colored glasses, and "water", more specifically OH (hydroxyl) radicals, which exhibit absorption bands in the near infrared. We examine these two types of impurities in turn.

a) Transition metal ions

Most transition metal ions of the ion series (Fe,Cu,Ni,Cr,V, Mn,Co), when inserted in glass hosts, exhibit more or less intense and broad absorption bands, located in various portions of the spectrum, from the very near infrared ($\simeq 1\,\mu m$) to the near U.V ($\simeq 350$ nm). Responsible spectral transitions take place between various electronic states of the ions. In glasses, these absorption lines are broadened mainly through fluctuations of the local environment field, from ion to ion (inhomogeneous broadening), so that *line-widths are indeed generally quite broad* (from a few 10 to a few 100 nm). In addition, a particular metal can give rise to absorption bands quite different in position and shape, depending upon the *state of valence* (vz of oxidation) it may assume. For instance, Fe^{2+} has a very broad band centered at around 1100 nm, whereas Fe^{3+} has a somewhat narrower band centered at around 350-400 nm : The relative concentration of different valence states, for a particular element, are strongly dependent upon the type of glass host and the preparation process and may also be influenced by the (deliberate or not) addition of other dopants that may displace redox equilibria. Also, relative absorption intensities and bandwidths , for a particular *ion*, may vary from glass to glass. It is therefore easy to understand that a particular glass host containing several transition metal impurities will exhibit very complex absorption spectra.

Figure 1 shows typical spectra of the most important transition ion impurities, obtained in a soda-lime silicate glass, (after [12]). Obviously, the main offenders appear to be Cu^{2+}, Fe^{2+}, Ni^{2+}, for the wavelength range of interest (800 - 1100 nm). Spectra shown on figure 1 are representative of a particular glass composition prepared through a *particular process*. For the same composition, other preparation processes lead to different relative absorption intensities per unit impurity *atom* concentration, as, in particular, relative concentrations of different valence states may vary when changing the process.

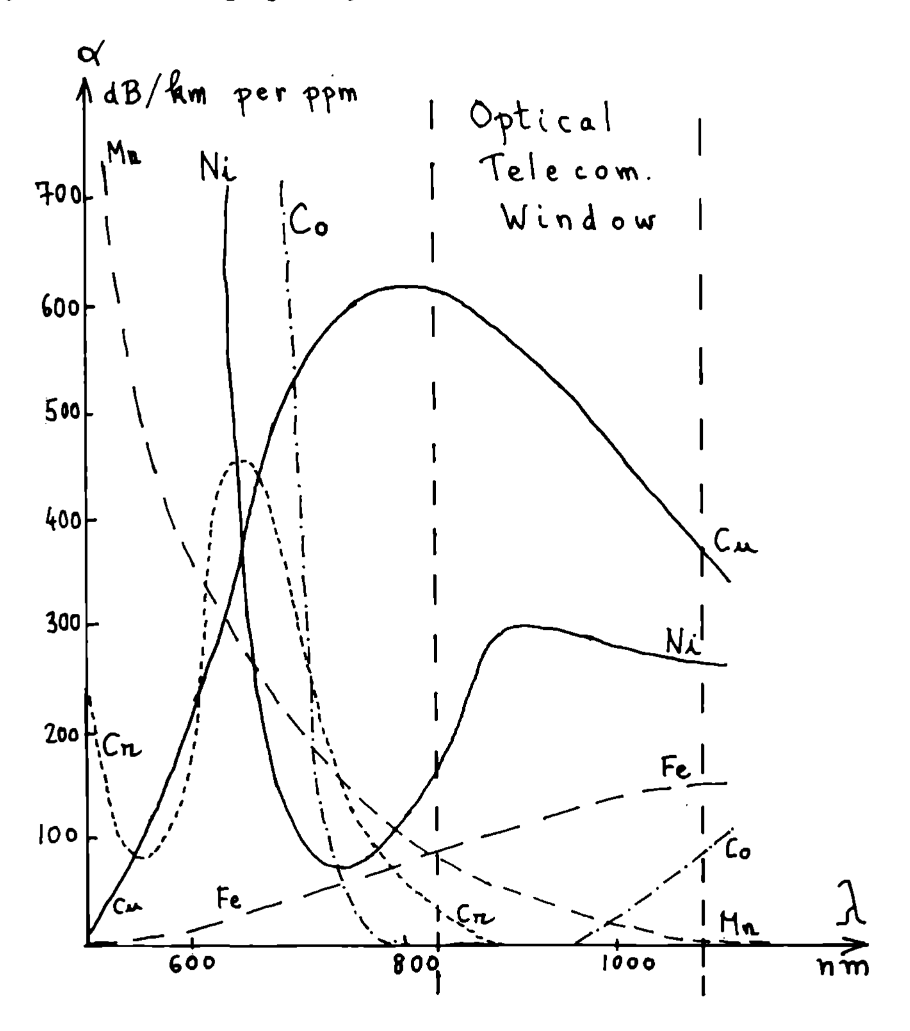


Figure 1 :
Absorption loss for various impurities as a function of wavelength.

Excellent discussions of optical absorption of transition elements in pure silica glass, containing also data on soda-lime silicate glass, are given by P.C. SCHULTZ [13] and by R.D. MAURER [17]. Tables II, III and IV, give typical examples of existing litterature data on soda-lime silicate glasses (II), actual experimental data on soda-lime glass (III) and fused silica (IV).

Ion	λ_{max} (nm)	Maximum ion concentration for absorption $\alpha_a \lesssim 20$dB/km at λ_{max}	λ =800 nm
Cu^{2+}	800	9 ppb (ion)	9ppb (ion)
Fe^{2+}	1100	8 "	15 "
ni^{2+}	650	4 "	26 "
V^{3+}	475	18 "	36
Cr^{3+}	675	8 "	83 "
Mn^{3+}	500	18 "	1800 "

Table II :
Absorption due to transition ions in a soda-lime silicate glass (SiO_2 70% - Na_2O 20% - CaO IO%)
(After R.D. MAURER [17])

Note : *Concentrations are in IO^{-9} (ions) units (or ppb (ion))*

Atom	Concentration (ppb (at))
Fe	$\simeq$ 200
Cu	< 10
Ni	< 10
Cr	< 50
Mn	< 10

Measured absorption at λ = 1060 nm : $\alpha_a \sim 27$ dB/km

Table III :
Absorption due to transition elements in a soada-lime silicate glass (SiO_2 70-75% - Na_2O 20-24% - CaO 3-9%)
(After C.E.E. STEWART et al [9])

Note : *Concentration are in IO^{-9} 'atoms/units (or ppb (at)) and relate to analytical measurements .*

Atom	Maximum *atom* concentration for absorption α_a<20 dB/km at λ= 88 nm	Attributed to ion :
Cu	2140 ppb (at)	Cu^{2+}
Fe	425 "	Fe^{2+}
Ni	712 "	Ni^{2+}
V	19 "	V^{3+}
Cr	33 "	Cr^{3+}
Mn	833 "	Mn^{3+}

Table IV :
Absorption due to transition elements in a fused silica glass (after P.C. SCHULTZ ([13]))
Note :
Concentrations are in IO^{-9} (atoms/units (or ppb (at))) and relate to analytical measurements

These data clearly illustrate how composition and processing may affect absorption spectra through changes in relative concentrations of different valence states of each responsible impurity. For example, as mentioned by P.C. SCHULTZ [13] it is most probable that, in fused silica glass made by flame hydrolysis of silicon compounds, Fe is essentially present in its oxidized Fe^{3+} state, and Cu in its reduced Cu^{+} state, so that this particular type of glass may tolerate comparatively large amounts of Fe and Cu (see Table IV) with no dramatic resulting increase in absorption at 800 nm. On the other hand, most published results on soda-lime silicate as well as alkali-borosilicate glasses [14] show that, with some variations due to the above-mentioned reasons, transition element concentrations and especially Fe et Cu, should be maintained at levels below 10^{-6} or even 10^{-7}, if contribution of impurity absorption to loss is to be less than 20 dB/km, in the 800 nm region.

It should be noted that this level of purity makes those glasses comparable to, or even better than the purest semiconductor materials. Those considerations provide some guide-lines to people who design and manufacture materials for low-loss optical waveguides. To sum up :

- if conventional composition glasses (silicates, alkali silicates, possibly lead glasses etc) are chosen, one should start with extremely pure raw materials (silica powder, alkali and earth-alkali salts) and try to maintain this purity throughout all processing stages, from mixing of powders up to fining and annealing. At the same time, careful attention should be given to all possibilities (particular choice of composition, controlled addition of oxidizing or reducing dopants, control of process

temperatures etc) of finally obtaining undesired transition element impurities in the least damaging valence state :i.e., for instance Fe^{3+} and Cu^{+}.

- in the case of "high-silica" glasses, most known techniques lead to directly produce the fiber- "preform" through oxidization (vz-hydrolysis) of silicon (and index-controlling dopant) compounds, mostly hydrides and halides [2][15]. In that case, the purity of the starting compounds is still important. However, it seems that the implied chemical reactions leave the most undesired impurities in the least damaging states. This probably partly explains why, to date, high-silica optical waveguides appear as "winners", as far as optical loss alone is concerned.

b) Hydroxyl radicals

All known glasses contain a certain amount of water, more precisely of OH groups. The fundamental vibrational band of OH gives rise to an infrared absorption line, which lies between 2.7 μm (fused silica) and 3 to 3.5 μm(other glasses. Detailed band width and shape depend of course, here again, of the composition. To this infrared line correspond successive harmonic lines. Although much weaker, some of those harmonics are of great importance,when one looks for ultra-low loss material at around 800 nm. Most important harmonic, in that respect, is the third one, i.e. at $\lambda \sim 925$ nm for fused silica. There also exist combination tones between OH - harmonics and Si-O vibrations, resulting in additional absorption lines in that region. Figure 2 shows typical absorption spectra related to those bands. Detailed analysis of those features has been made by D.B. KECK et al [3]. In fact, these absorption bands seem, to date, the only remaining limiting factor, besides intrinsic absorption and or scattering, for ultra low-loss "high-silica" waveguides.

As a rule of thumb, it is estimated that contribution of OH to the absorption, at the peak of the 3rd harmonic line (~ 925 nm) is of the order of :1.25 dB/km/10^{-6}(OH). High-silica low-loss fibers announced during the last two years still contained up to 50 to 80×10^{-6} OH. More recently however, further progress has been announced [2] on the removal of OH, with 3rd harmonic peak absorption coefficients as low as 4dB/km.

B) Scattering losses

I) Intrinsic scattering

Glasses are inherently light-scattering materials as they, at least, incorporate frozen-in density (i.e, refractive index) fluctuations. Those inevitable density fluctuations are related to what is known as "intrinsic scattering" which sets an ultimate limit, for a given material. The intrinsic scattering coefficient may be described [11] as :

$$\alpha_{Si} = \frac{8}{3} \frac{\pi^3}{\lambda^4} (n^8 p^2) (kT_f) \beta_T \qquad (2)$$

where λ is the optical wavelength.
n : the refractive index p : the photoelastic coefficient
β_T: the material compressibility
T_f: is the "fictive temperature" we already used for describing intrinsic absorption.
Quick inspection of (2) indicates the following :

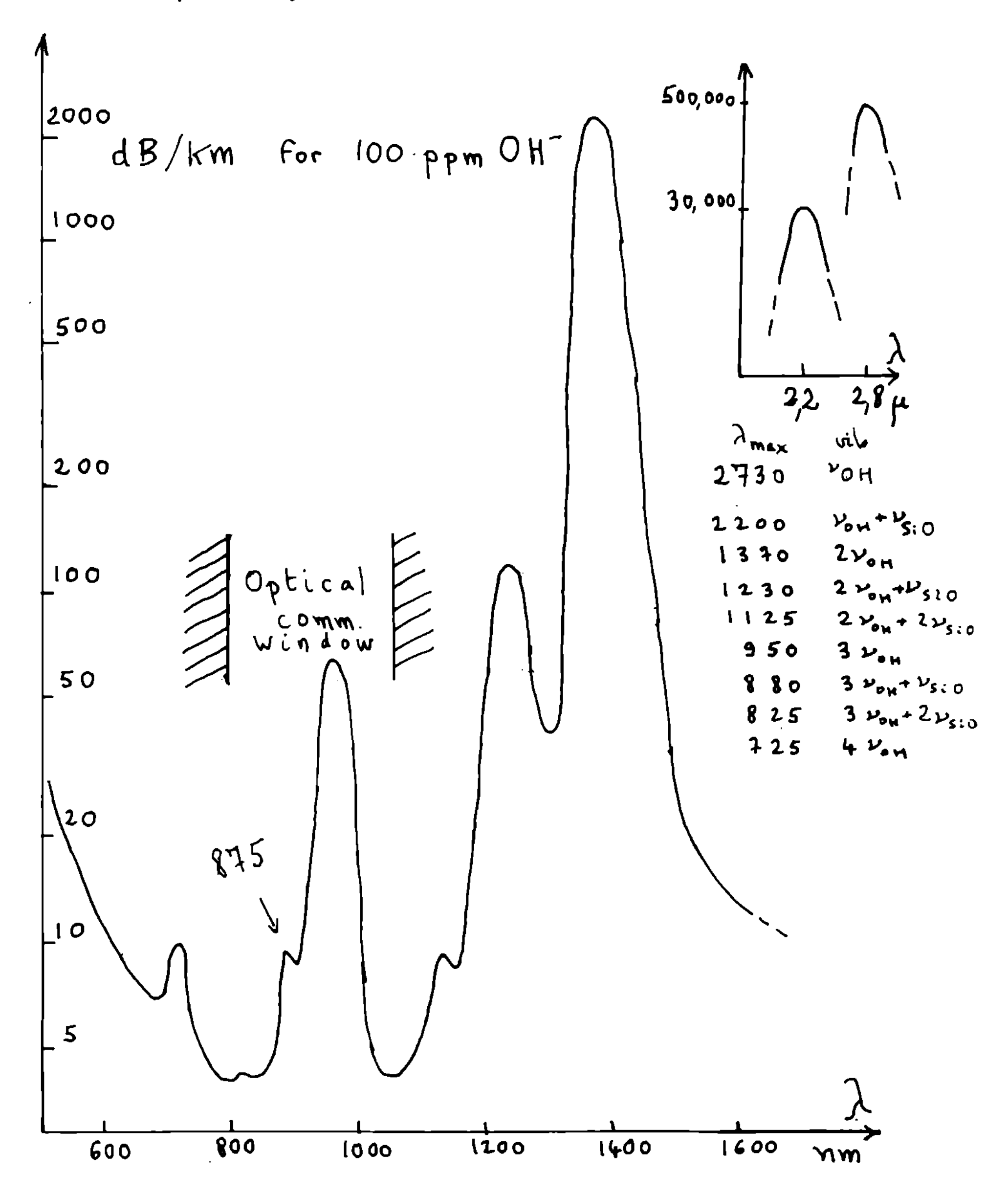


Figure 2 :
Influence of OH⁻ ions in silica.

• intrinsic absorption follows the famous Rayleigh-type λ^{-4} dependence.

• materials with comparatively high refractive index, photoelastic coefficient and compressibility will exhibit larger scattering losses.

• materials processed at comparatively high temperatures (melting, fining annealing) should also have larger losses. With respect to the last feature, one is led to conclude that the glasses mostly used for low-loss waveguides, range in quality from alkali borosilicates ($T_f \simeq 500$ K) through soda-lime silicates ($T_f \simeq 800$K) to fused silica ($T_f \simeq 1700$ K), in that order. That comparison may hold for "conventional" composition glasses, like the first two. However it is found that fused silica of best quality has lower scattering loss than most other glasses. This simply means that, for more complex glass compositions, other scattering mechanisms, are predominant, like, for instance, concentration fluctuation scattering.

II) *Concentration fluctuation scattering*

In a glass made up of several compounds, the inevitable statistical fluctuations in concentration of the different components result in a scattering loss coefficient of the type :

$$\alpha_{S,c} \propto \frac{16\pi^3 n}{3\lambda^4} \sum_i \left(\frac{\partial n}{\partial C_i}\right)^2 \overline{\Delta c_i^2} \qquad (3)$$

where c_i is the concentration of the ith component, there again, the λ-dependence is of the Rayleigh-type. Also, glasses with high refractive index will exhibit high $\alpha_{S,c}$, as $\frac{\partial n}{\partial c}$ increases with n. This may rule out, for instance, lead silicate glasses, that have been considered for making optical wave-guides [16] .

On top of the above mentioned microscopic scattering processes, glasses may show macroscopic defects, like large-scale index inhomogenities, bubbles, cracks , localized crystals etc..... that lead to scattering (usually not with a Rayleigh wavelength - dependence). Most of those defects, however, may be eliminated through careful processing technology. Removing those defects is, at any rate, a very first requirement for making a waveguide-grade glass material.

Based on intrinsic and concentration fluctuation scattering considerations,comparison between pure silica material and soda-lime silicate, for ultimate loss limits, appears, here again, to be in favor of silica , as shown on Table V (after D.A. PINNOW et al [11]) .

Type of glass	$\alpha_S = \alpha_{Si} + \alpha_{Sc}$ ($\lambda \simeq 900$ nm)
SiO_2	$\simeq$ 1 dB/km
SiO_2-Na_2O-Ca-O	$\simeq$3 to 4 dB/km

TABLE V

Type of glass material	Lowest reported loss Bulkglass	Lowest reported loss Fiber	Estimated absorption	Estimated scattering	Manufacturing	Wave-guide making
"High-silica"	(preform)→	< 2dB/km	≃1dB/km	≃1dB/km	Direct chemical synthesis of preform High temperature (≃ 1800°C)	Doping with index-raising (GeO_2, P_2O_5, $A\ell_2O_3$) or lowering (B_2O_3) oxides
Alkali borosilicates	?	14dB/km	≃10dB/km	≃4dB/km	Glass melting Low temperature (≃ 1300°C)	Double crucible drawing with alkali-ion exchange for index profile
Soda-lime	≃20dB/km	≃25-30 dB/km	?	?	Glass melting Intermediate temperature (≃ 1500°C)	Double crucible drawing Variation in composition for index change

TABLE VI

<u>Note</u> : all loss figures are for $\lambda \simeq$ 800 - 850 nm

III. COMPARISON OF PRESENT-DAY EXPERIMENTAL RESULTS

Considerable progress has taken place in the last 3 to 4 years, leading to rapidly improved loss performance for different types of glass material. One has to emphasize however that, after a period of "scanning through" many different possibilities, the choice seems now to be restricted to a limited number of glass composition types, that seem to best satisfy trade-offs between low-loss performance and other requirements, as evoked in I. We present (Table VI) a condensed comparison between these different types of materials, based on existing litterature (September 1974).

CONCLUSIONS.

Progress in the area of materials for low loss optical waveguides have been spectacular over the last 3 to 4 years. We can sum up the present day situation, as comment to the comparison outlined on Table VI, in the following way :

High-silica glass optical waveguides, presently available in several laboratories, are , with respect to loss performance alone, far beyond the best values that most people expected to reach, only 4 years ago, i.e. $\simeq$ 20 dB/km. The lowest loss figures reported for those waveguides are very close to the best estimates for ultimate limitations. However, it is known that many problems related to manufacturing, and handling of those waveguides in view of fabricating practical cables, have yet to be solved. In this respect, high-silica waveguides may suffer from mechanical problems, like a lesser resistance to fracture. Also, they might also be limited, for some specific applications, in the range of core/cladding inddex-differences (viz. numerical aperture).

Alkali-borosilicates and soda lime silicates are making steady progress, although perhaps at a slower pace than high-silica materials. However, there is no reason to believe that they have reached their ultimate limitations. In the long run, they may offer additional advantages with respect to fabrication cost, and versatility in ottaining the desired index-profiles. It is most probable that the first operationnal kilometric-distance optical communication systems will be heavily based on high-silica waveguides, However research is actively going on the other types of materials. This may eventually result in a situation where different types of waveguides could be used, depending upon specific applications.

REFERENCES

[1] K.C. KAO et al . Proc. I.E.E.E. Vol 113 , July 1966 ,pp. 1151-1158

[2] W.G. FRENCH et al "Optical waveguides with very low losses" Bell. Syst. Tech. J. May June 1974 , pp. 951-954

[3] D.B. KECK et al "On the ultimate lower limit of attenuation in glass optical waveguides"
Appl. Phys. Lett. Vol 22, n°7 , 1 April 1973, pp.307-309.

[4] T. UCHIDA "Optical Tele-communication Systems Using Fibers"
Proceedings of this conference. p. 341.

[5] J. STONE "Optical Transmission in Liquid-Core Quarts Fibers"
Appl. Phys. Lett. Vol 20, April 1972 , pp. 239-240.

[6] W.A. GAMBLING et al. "Gigahertz Bandwidths in Multimode, Liquid Core Optical Fibre Waveguide"
Opt. Commun. Vol 6 , Dec. 1972, pp. 317-322.

[7] M. TREHEUX "Propagation dans les Fibres Optiques"
Proceedings of this conference. p. 361

[8] R.D. MAURER , P.C. SCHULTZ "Fused Silica Optical Waveguide"
U.S. Patent 3 659915 , Assigned to Corning Glass Works
also :
L.G. VAN UITERT et al. Mater. Res. Bull. 8, 469 (1973)

[9] A.D. PEARSON, W.G. FRENCH, Bell Lab Rec. 50 [4] p.103 1972
and :
C.E.E. STEWART, P.W. BLACK, Electronics Letters, 7 May 1974
Vol 10, n°5 pp. 53-54.

[10] T. UCHIDA et al. I.E.E.E. J. Quantum Electronics Vol QE 6
pp. 606 - 612 Oct. 1970.

[11] D.A. PINNOW et al. "Fundamental Optical Attenuation Limits"
Appl. Phys. Lett. Vol 22, n°10, 15 May 1973 pp. 527-529.

[12] G.R. NEWNS et al. "Absorption losses in Glasses and Glass Fibre Waveguides".
Opto-electronics 5 1973. pp. 289-296.

[13] P.C. SCHULTZ "Optical Absorption of the Transition Elements in Vitreous Silica"
J. Amer. Ceram. Soc. Vol 57, n°7 July 1974 pp.309-313.

[14] G.R. NEWNS et al "Low-loss Glass for Optical Transmission",
Electronics Lett. Vol 10 n°10 , 16 May 1974 p. 201.

[15] D.B. KECK, P.C. SCHULTZ, U.S. Patent 3711 262

[16] A. JACOBSEN et al "Absorption and Scattering Losses in Glasses and Fibers for Light Guidance"
J. Am.Ceram. Soc. Vol 54, April 1971, pp. 186-187.
The following two references are review articles on glass material problems :

[17] R.D. MAURER , " Glass Fibers for Optical Communications"
Proc I.E.E.E. , Vol 61, n°4 April 1973 pp.452-462

[18] M. PASSARET "Les Matériaux pour Fibres Optiques et leur Caractéristiques"
CNET. Ann.Telc. 29 ,n°5-6 Mai-Juin 1974 pp.179-188.
The following reference is a review of all subjects pertaining to fiber optical communications, as of August 1973.

[19] S.E. MILLER et al "Research Toward Optical-Fiber Transmission Systems Parts I,II"
Proc. I.E.E.E. , Vol 61 , n°12 Dec 1973 , pp. 1703-1751
This article contains 260 references.

PERSPECTIVES EN INFORMATIQUE

F.H. Raymond
Thomson C.S.F.
Boulevard Haussman Paris

▪ I. INTRODUCTION

Posons-nous les questions suivantes :

- L'informatique attend-elle quelque chose de la photonique ?
- La photonique peut-elle révolutionner l'informatique ?
- A quelles conditions ?

A. L'informatique sollicite-t-elle la photonique ?

1) *La réponse est NON.*
Ce non peut être nuancé, puisque les transmissions sur fibres optiques intéressent les informaticiens ; mais ils ne sont pas préoccupés des progrès à accomplir pour les introduire dans leurs systèmes. En bref, ils ne sont pas moteurs. NON parce que les informaticiens ont à digérer le L.S.I. et ses développements. NON, car ils savent que les mémoires à disques, tambours, floppydiscs peuvent être perfectionnés sans augmentation de coût, du moins celle-ci sera absorbée par l'accroissement des débouchés et l'accroissement de la durée de vie de ces produits.

2) *L.S.I. et Floppy-disc .*
Voilà deux sujets qui détournent , avec raison, les informaticiens de la photonique. Pour un temps.

En effet, il est parfaitement concevable de construire pour moins de 1.000 dollars un "processeur" (miniordinateur spécialisé ou non) dans le volume d'un paquet de cigarettes, dont les caractéristiques sont comparables aux ordinateurs des années 50. Quelques pas de plus et dans le même volume pourra être réalisé un "ordinateur des années 70 ...".

Isolément, il est inapte à résoudre des problèmes.

L'idée de fédérer un certain nombre de "processeurs" , disons des robots, ainsi réalisés, est naturelle.

Ils sont fédérés avec des fichiers (protégés et aux accès contrôlés) et avec des moyens de communication et de dialogue avec les utilisateurs, ou l'environnement. D'où un système de communications programmées.

3) *Conséquence* .

Il se pose à l'informatique un problème de conception qui doit intégrer le logiciel et le matériel et, c'est une opinion, il bouscule les habitudes léguées par le développement industriel de l'informatique.

Le recours aux moyens technologiques coutumiers aux informaticiens pour concevoir un système de communications programmées conduit à accroître le coût d'une mission remplie à un niveau de complexité plus bas par les "bus" familiers aux informaticiens.

D'où la réponse à la deuxième question.

B. La photonique peut-elle révolutionner l'informatique ?

1) *Un système informatique (SI) apparaîtra comme étant :*
un système de communications programmables entre :

○● des mémoires
○● des "robots" (processeurs)
○● l'environnement (utilisateurs)

Oui la photonique peut révolutionner l'informatique :

● au niveau de la conception du système de communications et des moyens mis en oeuvre

●● au niveau des organes de dialogues avec le S.I.

On notera la différence de nature entre ● et ●● .

Les conférences de ce Colloque laissent percer des espoirs et nous devons être optimistes.

Si la proposition évoquée , contenant la réponse OUI la deuxième question posée au début, se réalise, l'effet d'entraînement pourra conduire :

○● à faire preuve d'imagination, à nouveau (si la Physique le permet !) pour les mémoires

○● puis pour les processeurs.

Retenons : le plus urgent sera ● et ●●

2) *Qu'est-ce que l'informatique peut attendre de la photonique ?*

a) Possibilité de transmettre une "image" par un faisceau lumineux

C'est-à-dire de transmettre en parallèle l'information contenue dans une image formée de "points" noirs ou blancs. La quantité d'informations ainsi transmises est sans commune mesure avec celle que n fils peuvent transporter simultanément , n étant nécesaisairement très inférieur aux nombres de points contenus ou discernables dans une image à transmettre.

b) Non interaction entre faisceaux lumineux permettant des organisations de systèmes dans lesquelles des faisceaux se croiseraient ; cette propriété n'est pas fondamentale, c'est une souplesse supplémentaire.

c) Possiblités potentielles d'utiliser des interactions photons-état solide pour :

1°) Enregistrer l'information contenue dans un faisceau de photons incidents : effet mémoire si l'état antérieur du solide bombardé par les photons est sans action sur l'état résultant des photons incidents ;

2°) Composer et enregistrer de l'information. Si, au contraire du cas précédent , l'état résultant des photons incidents dépend de ceux-ci et de l'état antérieur de la cible, à l'effet mémoire est associé un effet combinatoire qui peut s'exprimer par une "opération binaire" effectuée en parallèle sur l'information contenue dans l'état antérieur du solide et sur celle véhiculée par les photons ; masquage et sélection en sont deux applications.

d) Possibilités potentielles d'utiliser des interactions photons-état solide pour "moduler" un faisceau de photons en fonction de l'état du solide éclairé :

Si le "faisceau sortant" transporte l'information contenue dans le solide , on a l'effet lecture-mémoire en parallèle.

Si le faisceau sortant transporte le "produit" de l'information contenue dans le faisceau incident par celle contenue dans la cible, l'effet lecture est accompagné d'un effet combinatoire : la fonction combinatoire peut être le but poursuivi.

e) Dans tous les cas, les opérations évoquées sont effectuées en parallèle.

f) La possibilité de choisir des chemins de transport d'informations par une commutation globale : la déviation d'un faisceau pourrait coûter par bit transporté et nombre de cibles à atteindre par unité de temps considérablement moins que les solutions électroniques accomplissant la même fonction.

C. Interfaces

a) Dans un premier temps, les propriétés "parallèlisme" et "densité de stockage" conduisent à l'idée de "mémoire de masse" autour desquelles s'organiseraient les systèmes informatiques électroniques.

Des problèmes d'interface se posent :

a) Transformer des "paquets de bits" page- en image : c'est le concept d'imageur.

A son sujet, on peut présenter quelques remarques :

1°) L'information véhiculée par des signaux électriques l'atteint séquentiellement, même si un certain niveau de parallèlisme est mis en oeuvre : celui-ci ne dépassera pas la longueur des mots les plus longs manipulés par le système informatique (de 32 à quelques multiples de 32 augmentés des bits autocorrecteurs et détecteurs d'erreurs). Le débit incident à imager ne semble pas devoir dépasser le gigabits par seconde , à terme.

2°) Une page est transférée au milieu d'enregistrement par le faisceau de photons lorsqu'elle a été composée entièrement.

3°) Une page peut être composée morceau par morceau et simultanément transférée au milieu d'enregistrement.

En bref, le choix du rapport séquentiel -parallélisme au niveau de la fonction imageur est ouvert dans le cadre d'une page et peut poser le problème de l'interface sous la forme de la dualité temps - fractionnement de la page à "imager".

c) Transformer une image de page mémorisée en une page de signaux électriques.
Les même remarques peuvent être faites.

d) Il semble que les problèmes d'interface et les limitations dues à l'optique dans le cadre des projets concevables à ce moment -ayant beaucoup d'incertitudes par ailleurs- qui conditionnent le coût d'un système mémoire éliminent :
- Les dits projets comme concurrents à terme des disques magnétiques.
- Imposent l'objectif d'une grande capacité -10^{10} à 10^{12} bits - et un grand débit d'informations - plusieurs gigabits/sec.

Ceci conduit à ne pouvoir établir des projets de mémoire réalisables à moyen terme : l'idée du "bidimensionnel" est en fait l'élimination des projets les plus simples , dont pourtant les fonctions fondamentales n'ont pas été prouvées être réalisables industriellement : milieu d'enregistrement, imageur, interface de sortie.

D. Les mémoires dominent l'informatique

C'est une évidence, mais historiquement, les choses se sont passées différemment : les grandes capacités de stockage furent -et sont encore- offertes par les bandes et les disques magnétiques (les cartes et bandes perforées seront remplacées par des floppy-disc).
Dès lors que le prix du bit stocké sera inférieur disons à 10^{-4} Francs, les opérations à effectuer sur des informations peuvent l'être à l'aide de tables ce qui peut- et en principe c'est possible- faire disparaître les "circuits logiques et de calcul" : dans ces conditions, la transmission des informations au sein d'un système informatique reste le problème fondamental : l'optique intégrée est une voie qui va au delà des solutions imaginées à l'aide des circuits L.S.I.

Inversement, il pourrait sembler plus raisonnable d'entreprendre des recherches sur l'application de l'optique intégrée à l'organisation de "circuits logiques et de calcul ". Bien que toute idée dans cette voie ne doive pas être éliminée à priori, le poids de l'interface fait craindre que ceci ne soit intéressant industriellement que si une masse considérable de traitement doive être accomplie. Alors, la complexité croît très vite et l'issue semble bien être dans la recherche d'une certaine régularité : à l'extrême, la plus grande régularité est obtenue avec les mémoires. D'ailleurs, je suis convaincu que l'impact des semi-conducteurs dans le domaine des mémoires conduira à cette évolution : il est plus logique de gâcher de la technologie produite en série et produisant des produits "testables" que d'augmenter la variété de circuits compliqués difficilement testables.

Mais au niveau actuel des L.S.I. la preuve du contraire est valide.

E. A quelles conditions la photonique révolutionnera-t-elle l'informatique ?

La réalisation de quantités considérables d'intra et d'interconnexions domine, aux plans conceptuel et industriel, l'informatique : d'où des procédés collectifs de réalisation.
La photonique subira cette contrainte d'où l'optique intégrée.

Les coûts évalués , au niveau d'un projet, devront être d'un ordre de grandeur moindre que ceux des solutions purement électroniques utilisées à l'époque ou susceptibles de l'être (cas d'un sytème complexe d'inter-communications entre robots et organes de dialogues).
Le gain en fiabilité et maintenabilité devra être significatif afin qu'un changement de technologie améliore le service rendu par les produits informatiques et procure des économies aux parties prenantes : utilisateurs et constructeurs.

LA PHOTONIQUE, TECHNIQUE DE DEMAIN

P. Aigrain
Thomson CSF
Bd Hausmann, Paris
France

Les raisons évidentes pour lesquelles j'aurais aussi bien pu donner l'exposé d'introduction que l'exposé de clôture, sont que contrairement à tous les autres conférenciers ici, qui avaient quelque chose à dire de précis , qui présentaient des résultats techniques nouveaux, étant actuellement actifs dans le domaine, je n'y connais pas grand chose. Par conséquent tout ce que l'on peut me demander, c'est de faire de la philosophie.

Vous vous souvenez de la définition de la différence entre l'ingénieur, le physicien, le mathématicien et le philosophe. L'ingénieur lui, a un problème, mais il ne connait pas toujours les lois physiques qui s'appliquent à son problème, il est obligé de demander au physicien de lui exprimer ces lois physiques. Malheureusement le physicien après cela ne sait pas résoudre les équations qui en résultent et par conséquent il s'adresse au mathématicien qui lui, résoud les équations. Et puis tout revient au point de départ, et le problème est résolu. A ce moment-là arrive le philosophe. Il ne connait pas le problème, il ne connait pas les lois physiques qui le régissent, il ne connait pas les manières de résoudre les équations mathématiques, mais il en parle. Voilà ce que l'on m'a demandé de faire....

Le titre est : "LA PHOTONIQUE, TECHNIQUE DE DEMAIN". Ce n'est pas moi qui ait inventé ce titre, ni même Monsieur Balkanski, mais je pense qu'il a eu raison de le choisir. Il implique un premier point: que très évidemment la photonique n'est pas encore tout à fait la technique d'aujourd'hui. Si l'on veut imaginer un petit peu son avenir, je pense qu'il pourrait être utile de se livrer à un rêve humoristique qui serait basé sur l'idée du monde à l'envers; de ce qui se serait produit si les techniques s'étaient développées dans un ordre différent de celui que nous avons connu.

Imaginons que nous soyons en train de nous réunir, ici à Cadarache, à un instant que je ne préciserai pas, dans un beau château qui serait la maison d'hôtes, non pas du Commissariat à l'Energie Atomique pour rester cohérent avec le thème de mon rêve, mais par exemple du Commissariat à l'Energie Charbonnière créé par le gouvernement français. Celui-ci aurait commencé à s'inquiéter depuis une quinzaine d'années et un peu plus même de l'épuisement des ressources d'uranium, nous amenant à manquer d'énergie. Par conséquent il allait falloir considérer une source certes bien plus difficile à exploiter parce que mal connue et qui s'appelait le charbon. Alors on aurait créé un Commissariat à l'Energie Charbonnière, et des chercheurs seraient en train de travailler sur les nombreux problèmes que pose l'utilisation de ce combustible, car après tout, il faut le faire réagir avec l'air, et vous comprenez, il y a de l'oxygène pouvant faire de l'oxyde de carbone, il y a de l'azote pouvant donner du cyanogène qui est violemment toxique. Vous imaginez les problèmes, et par conséquent ce Commissariat étudierait cette nouvelle technique. Et puis nous, nous serions ici pour étudier une nouvelle technique qui serait en train de paraître et qui s'appellerait l'électronique compte tenu de ce que dans un passé déjà relativement ancien, les méthodes de traitement de l'information auraient fait d'énormes progrès, qu'elles seraient largement diffusées, mais qu'elles seraient entièrement fondées sur l'utilisation de méthodes optiques.

Et alors brusquement des progrès qui auraient été réalisés récemment en matière de matériaux et en matière de technologie de mise en oeuvre, nous ouvriraient apparemment une porte nouvelle, qui serait celle de l'électronique. Et l'on se réunirait ici pour dire quel formidable progrès constitue l'électronique. Vous vous rendez compte de ce que l'on peut faire avec cela. Si vous diminuez les dimensions jusqu'à l'ordre du micron ou en dessous vous imaginez la densité d'informations que vous arrivez à mettre. Vous n'êtes plus limités par les longueurs d'onde. Vous pouvez maintenant réduire les dimensions encore bien plus loin. Vous n'avez plus ces limitations à la densité de traitement de l'information qui jusqu'à présent nous ont gênés dans les systèmes optiques. Nous n'avons plus les problèmes de fiabilité , de dégradation des lasers que l'on a résolus, mais à quel prix, avec quelles difficultés. Et visiblement comme on ne connaitrait pas encore très bien quelles sont les difficultés non moins réelles et non moins difficiles qu'ont eues les transistors, on se dirait qu'avec cela on arrivera très probablement à résoudre les problèmes d'avenir. Et on serait probablement très excité et on demanderait à quelqu'un qui n'est plus dans le domaine depuis dix ans de présenter une conférence sur l'électronique, technique de demain.

Si je me suis amusé à donner cette explication, aujourd'hui pourquoi la Photonique ?, c'est pour que nous regardions les choses d'un petit peu plus près. En réalité, en quoi la photonique va-t'elle présenter de l'intérêt? Et là je crois qu'il y a un certain nombre de pièges dans lesquels il ne faut pas tomber. Ne croyons pas

en particulier que l'intérêt de la photonique tient dans les grandes densités de l'information de l'optique intégrée. C'est presque le contraire. C'est parce que lorsqu'on fait de l'optique intégrée on est obligé d'avoir des dimensions qui sont de l'ordre de grandeur de la longueur d'onde, ni plus petites, ni plus grandes - autrement cela ne se propage pas - que l'on a de grandes densités d'information. Mais si l'on utilise les mêmes dimensions dans des dispositifs purement électroniques, ce qui est par définition possible, puisque ces dimensions sont limitées par des technologies de mise en oeuvre, eh bien, on a les mêmes densités et on en a même de plus grandes. Mon ami Ostrowsky ne m'en voudra pas. Je lui dirai qu'après tout il nous a présenté par exemple de remarquables commutateurs tel que le Cobra. Le Cobra fait appel à des microplanches et d'ailleurs il le faut, on ne peut faire autrement. C'est en partie à cause des développements de l'électronique que l'on sait faire des microplanches, mais cela fait quand même 500 µ de long, parce que c'est cela la longueur d'interaction, et si cela fait 500 µ de long, c'est très grand par rapport à un commutateur qui pourrait être réalisé avec deux transistors et quelques éléments associés en électronique intégrée.

Par conséquent, je ne crois pas que l'intérêt de la photonique se tienne nécessairement dans le domaine des augmentations de densité, qui me semble beaucoup plus liées aux progrès des technologies qu'à l'utilisation d'un nouveau support d'informations. De même ne croyons pas nécessairement que l'on va travailler avec des niveaux de puissance notablement plus bas. En fait, pour transporter un bit d'information, le minimum c'est d'avoir un photon. On ne peut pas faire moins. Et encore lorsque l'on n'a qu'un photon, on transporte le bit avec une probabilité de le reconnaitre qui n'est pas fameuse. Alors si l'on veut avoir des probabilités d'erreur qui soient raisonnables (pas nécessairement aussi bonnes que les 10^{-11} que l'on a citées ce matin, parce qu'après tout la redondance ça existe et l'on ne fait pas 10^{-11} avec les mémoires à semiconducteurs aujourd'hui), il faut quand même des 10 à 15 photons, et tout à l'heure on citait que les meilleures diodes à avalanche nécessitaient des nanowatts et j'étais un peu impressionné que les gens n'aient pas trouvé cela remarquable. En effet un nanowatt en une nanoseconde, cela fait 10^{-18} joules, soit 6 eV et 6 eV cela ne fait pas beaucoup plus qu'un photon. Il ne faut pas se tromper, si vous êtes à une nanoseconde, vous ne pouvez pas espérer détecter beaucoup moins qu'un nanowatt, même avec le meilleur PM du monde. Il y a une limite quantique. Il est intéressant de savoir que cette limite tient à ce que les photons ont une énergie de l'ordre de 2 eV, mais qu'après tout tant que l'on travaillait dans le domaine des ondes centimétriques cette limite n'existait pas. Bien sûr il y a toujours le kT qui intervient, mais c'est une limite qui est plus basse. Théoriquement les limites d'énergies consommables sont meilleures en électronique qu'en photonique.

Par dessus tout, nous avons une dernière difficulté en photonique. Je vais vous le dire tout de suite, c'est cette

difficulté même qui à mon avis fait quand même de la photonique, la technologie de demain: c'est le problème de l'homogénéité des techniques à l'intérieur d'un système. Toutes les fois que l'on a un système technique, par exemple un système de traitement de l'information, un système de transmission de l'information, de commutation, etc... il se révèle qu'une partie importante des difficultés et des coûts sont associés aux transitions d'une technique à une autre, aux interfaces entre techniques. Exemple: le problème du microphone pour passer des signaux acoustiques aux signaux électriques. Il est évident que dans un monde où les techniques de transmission et de traitement de l'information sont actuellement électroniques, basées sur le mouvement de l'électron, et sur les ondes électromagnétiques non optiques, l'introduction de la photonique se heurte à une difficulté de compatibilité avec le reste du système. Alors est-ce que cela veut dire qu'il n'y a pas d'espoir, compte tenu des énormes difficultés qui restent à résoudre. Je pense en particulier aux énormes difficultés liées à la fiabilité des sources, car nous ne savons pas encore (nous avons tous de l'espoir) s'il sera réellement possible de faire un laser à semiconducteur ayant une durée de vie sûre de 100 000 heures. Personne ne le sait, cela apparait de plus en plus probable mais les limites intrinsèques de la dégradation ne sont pas encore connues, parce que nous n'avons pas encore de théorie là-dessus. Alors est-ce que cela vaut la peine? ma réponse est oui après avoir donné tous les arguments négatifs, c'est à dire OUI, cela vaut la peine, parce qu'il y a deux domaines dans lesquels je pense qu'à terme l'optique s'introduira incontestablement.

Le premier, on en a beaucoup parlé ici, c'est le problème des transmissions. Mais tous les arguments que j'ai donnés contre la photonique semblent s'appliquer à fond au cas des transmissions optiques, et l'on se demande vraiment pourquoi diable penser à des photons alors que nous savons transporter des ondes électromagnétiques par exemple dans des guides d'ondes millimétriques, avec des atténuations très faibles dans des largeurs de bandes considérables et à niveaux de puissance détectables qui sont meilleurs que ce que l'on peut faire en photonique. Eh bien, la raison essentielle et qui je crois imposera les transmissions optiques dans un grand nombre d'applications, c'est qu'il y a un facteur technique qui est meilleur que partout ailleurs, et qui est meilleur dans une telle proportion qu'il est impensable que cela ne joue pas: c'est le nombre de communications qu'on arrive à faire passer par unité de section. Or le nombre de cas dans le monde moderne où le coût de la section de passage est devenu une partie importante du coût du système de communications est considérable. Bien sûr l'un de ceux qui viennent tout de suite à l'esprit est le problème des lignes d'abonnés dans les grandes villes dans lesquelles la difficulté n'est même pas le coût de la tranchée pour aller mettre la ligne, quoiqu'elle représente 70% du coût total des lignes d'abonnés - car quand on fait de la transmission, on est quelquefois étonné de voir que tout l'ar-

gent part dans le génie civil - mais c'est qu'il n'y a plus de place, il n'y a plus de trou à faire parce qu'ils sont tous pleins. Il y a déjà là-dedans des canalisations d'eau ou de gaz, des cables téléphoniques dont certains ne sont peut-être même plus utilisés, mais difficiles à retirer , et tout cela s'entrecroise. Si l'on continue, il va falloir descendre dans des profondeurs de plus en plus grandes, on se trouve alors en général dans des nappes d'eau sale sous les grandes villes et par conséquent un problème quasi insoluble est déjà en train de se poser dans les villes les plus denses, New York en particulier et même Paris qui est aussi dense que New York et qui est la seule ville de France dans laquelle la densité téléphonique est à peu près correcte. Bizarrement comme vous le savez si la France ne se place pas très bien dans le monde au point de vue densité téléphonique, par contre Paris se place très bien dans le monde parmi les villes à forte densité téléphonique (76 téléphones pour 100 habitants) . Ce qui n'est pas si ridicule par rapport à Washington à 126 téléphones pour 100 habitants, ce qui laisse supposer que certains restent inutilisés à un moment donné. Dans ces grandes villes, le problème se pose de manière aigue. Il n'y aurait que cela, je ne suis pas sûr que cela conduirait à un très grand développement, mais il y énormément d'autres cas où à défaut d'une impossibilité technique de résoudre le problème par d'autres méthodes que les transmissions optiques, l'avantage qui résulte de la faiblesse de la section à travers laquelle on passe la communication est très important. Une fibre optique ne vaut pas un guide d'ondes millimétriques quand on y regarde de près: la largeur de bande ne sera pas meilleure compte tenu de la dispersion même en la compensant par toutes les astuces qui ont été proposées et l'atténuation est encore beaucoup plus forte même avec les fibres de plus faibles pertes. Mais le guide d'onde millimètriques a plusieurs centimètres de diamètre, et la fibre optique même avec le guainage mesurera moins d'un demi millimètre. Ceci ne veut pas dire que ce sera bon marché si on en juge par le prix des fibres Corning, mais cela peut baisser sérieusement dès qu'il y aura de la concurrence.

Les fibres ne seront pas nécessairement bon marché, mais en tout cas elles seront légères et faciles à poser. Or le coût de mise en place va jouer un rôle important dans une quantité de cas. Donc pour des raisons essentiellement de faible encombrement de la fibre , beaucoup plus qu'à cause de ses avantages en tant que moyen de transmission, je crois que les transmissions optiques s'imposeront, car le faible encombrement, c'est peut-être un petit peu moins glorieux à mentionner que des largeurs de bande ou des cadences d'informations, mais cela joue un rôle tout aussi important.

Et puis, il y a un deuxième cas où d'hors et déjà l'optique joue un grand rôle, seulement elle le joue en général à travers des moyens qui devraient nous faire rougir de honte, c'est le cas des présentations de données, car finalement à l'exception du téléphone acoustique où la sortie est bien en-

tendu acoustique, dans presque tous les autres cas, les traitements d'informations que nous faisons débouchent en dernière analyse sur la présentation d'images: d'images imprimées, d'images placées sur un tube cathodique, sur un écran, présentées de toutes les manières que vous voulez, mais finalement l'oeil est de loin parmi les sens que nous a donnés la nature celui qui nous permet de recevoir la plus grande quantité d'informations interprétables. L'oreille a apparemment à peu près autant de capacité, mais c'est beaucoup moins interprétable. De telle sorte que d'hors et déjà, si vous avez par exemple un système de traitement de l'information, ces informations sortent en dernière analyse sur des bouts de papier ou sur des écrans cathodiques. Seulement quand on regarde comment est équipé un imprimeur, je vous dis c'est à faire rougir de honte le technologue du milieu du XXe siècle. C'est vraiment une technologie du XIX siècle construite, il faut le reconnaitre, avec des moyens de fabrication qui eux utilisent à fond la technique du XX siècle, ce qui d'ailleurs a permis à de telles technologies périmées de survivre. C'est étonnant que le tube cathodique qui a été découvert en 1920 reste encore notre moyen essentiel de présentation d'images: On n'a pas trouvé mieux pour le moment. Mais ces méthodes ont toutes pour but en dernière analyse de nous présenter de la lumière à nos yeux. Et il nous semble évident que du point de vue de l'homogénéité des techniques, il devrait être plus simple de pouvoir faire la présentation de données dans des systèmes de traitement de l'information où l'information serait traitée sous forme optique plutôt que sous forme électrique. Je ne sais pas comment d'ailleurs, et je pense qu'il y a une certaine logique à dire que si la sortie finale est optique dans 90% des cas, il faudra qu'une bonne partie de ce qui est avant soit de préférence optique.

Pour ces deux raisons, nous avons là des systèmes qui sont je dirai optiques par utilité ou par nature. Eh bien ces systèmes vont aller plus facilement avec des moyens de traitement de l'information et de transmission qui seront eux aussi optiques. Voilà pourquoi je crois effectivement que la photonique est une technologie de demain.

Je ne vais pas entrer maintenant dans le détail des problèmes à résoudre parce qu'ils vous ont été largement exposés ici, et que personnellement j'ai trouvé fort intéressant que presque tous les conférenciers aient indiqué les difficultés qui restaient à résoudre aussi bien que les possibilités d'avenir. Trop fréquemment dans les colloques de ce genre on ne présente que les choses qui ont marché. Ici très heureusement on a aussi indiqué quelles étaient les difficultés, et le fait qu'elles n'étaient pas encore toutes controlées. Mais je pense que tout indique que ces difficultés pourraient être controlées au prix d'un effort technologique qui sera probablement d'un ordre de grandeur comparable à celui qui a été nécessaire pour résoudre les problèmes de l'électronique intégrée moderne, c'est à dire considérable, mais qui pourra

bénéficier du fait que les technologies de mise en oeuvre sont très analogues à celles de l'électronique intégrée moderne. Vous réaliserez une quantité d'éléments par les vieilles méthodes, avec les mêmes dispositifs, que cela aille de l'application des techniques de photogravure ou de masquage électronique aux techniques de diffusion ou d'implantation d'ions, et ceci devrait permettre de gagner sinon de l'argent du moins du temps. A condition qu'effectivement ces problèmes se résolvent, et à mon avis deux des plus importants et des plus difficiles, sont bien sûr ceux des fibres de réellement bonne qualité, de réellement faibles pertes et de mise en oeuvre facile et ceux des sources. Je crois que ceux des détecteurs sont en fait beaucoup plus proches d'être résolus. Ce qui ne veut pas dire que l'on ne puisse pas encore faire des progrès. Et puis du côté du traitement de l'information, il y a tout le problème de l'optique intégrée, du traitement des informations optiques, de la logique et des opérations analogiques à faire par des méthodes optiques avec peut-être des possibilités de nouvelles technologies. Je pense en particulier aux possibilités qu'offrent des fabrications de certains éléments par réplique pressée qui sont peu applicables au secteur de l'électronique habituelle mais qui pourraient trouver un emploi dans le secteur de l'optique intégrée avec d'énormes avantages car de toutes les techniques de fabrication que je connaisse c'est de loin la meilleur marché et cela dans une proportion considérable. Le fait que l'on puisse imprimer par exemple toute une lentille de Fresnel, qui est quand même un appareil optique compliqué, d'un seul coup de presse et que les précisions que l'on obtient soient étonnamment bonnes, les reliefs se répliquent à une centaine d'angströms près sur une presse de caractère absolument élémentaire, devrait fournir des possibilités que je mentionne parce que justement on n'en a pas beaucoup parlé ici, mais qui existent à mon avis réellement.

Si vous continuez votre effort, et je suis persuadé que vous le continuerez dans la mesure où les différents bailleurs de fonds continueront à vous financer (cela inclut bien sûr les administrations des différents pays et les directions générales des grandes entreprises: un certain nombre de gens qui ont toujours le défaut de dire non quand on demande de l'argent - je suis doublement placé pour le savoir-), eh bien, s'ils continuent à vous soutenir, je crois que demain, c'est à dire en 1990, la photonique aura pris une part importante du marché du traitement et de la transmission des informations. Demain donc, en 1990, la photonique sera la technologie d'aujourd'hui, et peut-être pourra-t'on refaire mon petit rêve de tout à l'heure et essayer de voir si par hasard on ne peut pas faire mieux avec la bonne vieille électronique.

I N D E X

Imprimerie GAUTHIER-VILLARS, 70, rue de Saint-Mandé, 93100 MONTREUIL
190803-75 — Dépôt légal : 1er trimestre 1975 Imprimé en France

Page d'auto-publicité

BALKANSKI - Photonique

DUNOD et GAUTHIER-VILLARS vous proposent :

Th. KAHAN et al.
Théorie des groupes en physique classique et quantique
Tome 1. Structures mathématiques et fondements quantiques
688 pages. 16 x 25. 2 fascicules brochés. Dunod.

Tome 2. Applications en physique classique
352 pages. 16 x 25. broché et relié. Dunod

Tome 3. Applications en physique quantique
300 pages. 16 x 25. broché et relié. Dunod

V. P. MASLOV
Théorie des perturbations et méthodes asymptotiques
Collection Etudes mathématiques

400 pages. 16 x 25. relié. Dunod

G. DUVAUT, J.L. LIONS
Les inéquations en mécanique et en physique
408 pages. 16 x 25. relié. Dunod

P. NOZIERES
Le problème à N corps. Propriétés générales des gaz de fermions
334 pages. 16 x 25. relié. Dunod

L. de BROGLIE, sa conception du monde physique
416 pages. 16 x 25. relié. Gauthier-Villars

A paraître fin 1975 aux Editions FLAMMARION
3rd *International Conference on light scattering in solids*
Campinas (Brésil), 28.7.75 - 3.8.75

Photonics

The word photonics (i.e. opto-electronics) describes, by analogy with electronics, all functions where the electron is replaced by the photon as the information vehicle, i.e. light emission, transmission, modulation, deflection, amplification and detection.

Recent developments of semi-conductor lasers, optical fibres and high sensitivity rapid detectors have opened the way for rapid and extensive utilization of photonics in telecommunications.

This book is a collection of papers by a number of physicists and engineers actively engaged in research in this field. The physical phenomena underlying each of the functions are discussed in relation to the materials which offer the best performance.

This collection is probably the first to cover the field of photonics in such depth. It gives the present status of development and we hope will also encourage further research. An invaluable tool for industrial and university laboratories looking for improved electro-optic materials.

La photonique (ou opto-électronique) couvre le domaine de recherche et d'applications dans lequel l'électron est remplacé par le photon comme support de l'information.

Les progrès récents en matière de lasers à semi-conducteurs, de fibres optiques de faible perte et de détecteurs de haute sensibilité conduisent à penser que la photonique trouvera de multiples applications dans les télécommunications.

Ce livre se compose d'exposés présentés au cours d'un colloque organisé par la DGRST sur les différents aspects de la photonique.

Les auteurs, physiciens ou ingénieurs, présentent les principes physiques utilisés dans cette nouvelle discipline, publiant sur les matériaux actuellement disponibles ainsi que sur les réalisations pratiques de systèmes, une mise au point très attendue.

C'est donc un ouvrage pionnier que ce "Photonics", premier du genre à fournir aux chercheurs travaillant dans les laboratoires industriels et universitaires, le moyen de rechercher des matériaux de meilleures performances et d'aboutir à des applications nouvelles.

ISBN 2-04-005023-X